



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 111490

TO: Mark Navarro  
Location: cm1/8a15  
Art Unit: 1645  
Tuesday, January 06, 2004

*Handwritten signature/initials inside a circle.*

Case Serial Number: 08/482785

From: Barb O'Bryen  
Location: Biotech-Chem Library  
CM1-6A05  
Phone: 308-4291

*Handwritten signature/initials.*

barbara.obryen@uspto.gov

### Search Notes

BEST AVAILABLE COPY

**This Page Blank (uspto)**



```

XX Claim 15; Page 89; 115pp; English.
PS
XX
CC A leader peptide (AAR88822) of Streptococcus pyogenes DNase B
CC can be used to facilitate prodn. Of recombinant DNase B (see
CC AAR88823) in transformed bacterial hosts, e.g. Escherichia coli.
CC The leader peptide can also be used for expression and prodn.
CC of other recombinant proteins in bacteria. The product is
CC excreted by the host into the culture medium and is easily
CC recovered.
XX
SQ Sequence 43 AA;

Query Match 100.0%; Score 203; DB 17; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.5e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43
DB 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43

RESULT 2
AAR88821
ID AAR88821 standard; Protein; 90 AA.
XX
AC AAR88821;
XX
DT 25-JUN-1996 (first entry)
XX
DE S. pyogenes DNase B partial sequence.
XX
KM DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;
KW diagnosis; Escherichia coli.
XX
OS Streptococcus pyogenes strain ATCC 14289.
XX
FH Key Location/Qualifiers
FT Peptide 1..43
FT /label= Sig_peptide

W09606174-A1.
XX
PD 29-FEB-1996.
XX
PF 18-AUG-1994; 94WO-US09450.
XX
PR 18-AUG-1994; 94WO-US09450.
XX
PA (BECI ) BECKMAN INSTR INC.
XX
PI Adams CW, Belei MC, Pang PPY;
XX
WPI; 1996-151377/15.
DR N-PSDB; AAT12773.
XX
PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant
PT prodn. of the enzyme in other bacteria, useful in immunoassays or
PT for treating cystic fibrosis
XX
PS Example 2; Fig 3; 115pp; English.
XX
CC A partial sequence (AAR88821) for Streptococcus pyogenes DNase B
CC is the product of DNA clone lambda 2-6 (AAT12773) isolated from a
CC S. pyogenes DNA library. A full-length sequence is given in
CC AAR88823. The insert of the DNase B-encoding DNA clone can be
CC transferred to expression plasmids for large-scale recombinant
CC DNase prodn. in E. coli or other bacterial hosts. It is used in
CC immunoassays to detect anti-DNase B antibodies in serum as a
CC marker of S. pyogenes infection, and is also useful as a vaccine
CC or for cystic fibrosis treatment via aerosol delivery.
XX
SQ Sequence 90 AA;

```

```

Query Match 100.0%; Score 203; DB 17; Length 90;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43
DB 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43

RESULT 3
AAR70700
ID AAR70700 standard; Protein; 91 AA.
XX
AC AAR70700;
XX
DT 25-MAR-2003 (updated)
DT 14-AUG-1995 (first entry)
XX
DE DNA-ase-B2 N-terminal fragment.
XX
KM DNA-ase-B2 N-terminal fragment; diagnostic; vaccine; cystic
KW fibrosis therapy.
XX
OS Streptococcus pyogenes.
XX
FN W09500650-A1.
XX
PD 05-JAN-1995.
XX
PF 18-MAY-1994; 94WO-US05626.
XX
PR 23-JUN-1993; 93US-0082845.
XX
PA (BECI ) BECKMAN INSTR INC.
XX
PI Adams CW, Belei MC, Pang PPY;
XX
WPI; 1995-052087/07.
DR N-PSDB; AAQ85036.
XX
PT New DNA encoding Streptococcus pyogenes DNase B - for diagnosing
PT S. pyogenes infection, also new promoter for expressing other
PT proteins
XX
PS Claim 3; Fig 3; 97pp; English.
XX
CC The sequence represents the N-terminal fragment of Streptococcus
CC pyogenes DNA-ase-B, which is a marker of S. pyogenes infection.
CC The protein may be used as a diagnostic agent or vaccine for S.
CC pyogenes, or may be used as an aerosol to treat excessive lung
CC viscosity, e.g. in cystic fibrosis.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 91 AA;

Query Match 100.0%; Score 203; DB 16; Length 91;
Best Local Similarity 100.0%; Pred. No. 3.7e-23;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43
DB 1 MNLLGSRVRSKKCRVLVFSWALVSATMAVTVTLTENTALAR 43

RESULT 4
AAR58702
ID AAR58702 standard; Protein; 271 AA.
XX
AC AAR58702;
XX
DT 25-MAR-2003 (updated)
DT 29-MAR-1995 (first entry)

```



```
XX DE Mitogenic factor associated with group A Streptococci.
XX XX mitogenic factor; microdetection; group A streptococci; spe;
KM erythrogenic toxin; streptococcal pyrogenic exotoxin; blastogens;
KM scarlet fever toxin; erythematous skin reaction; infectious disease;
KM delayed hypersensitivity; ss.
XX OS Streptococcus pyogenes.
XX PN EP613947-A2.
XX PD 07-SEP-1994.
XX PF 31-JAN-1994; 94EP-0101386.
XX PR 01-FEB-1993; 93JP-0037383.
XX PA (SHIO ) SHIONOGI & CO LTD.
XX PI Hara A, Hinuma Y, Igarashi H, Iwasaki M, Kishishita M,
PI Okumura K, Takeda Y, Yutendo T;
XX DR WPI; 1994-272994/34.
XX DR N-PSDB; AAQ71612.
XX CC New mitogenic factor gene from Streptococcus pyogenes - used to
PT develop probe, for the early diagnosis of infectious disease
PT caused by gp A streptococci
XX PS Claim 7; Page 12-13; 20pp; English.
XX CC AAR58702 shows a mitogenic factor which exhibits rabbit peripheral
CC blood lymphocyte mitogenicity and/or DNA hydrolysis activity. It is
CC strongly associated with group A Streptococci and the nucleotide
CC sequences can be used for the microdetection of the gene and provide
CC an early diagnosis of infectious disease caused by the bacteria.
CC (See also AAQ71613-26).
CC CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 271 AA;
SQ Query Match 100.0%; Score 203; DB 15; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLGSRVRFSSKCKLVKFSWALVSATMAVTTVLTENTALAR 43
Db 1 MNLGSRVRFSSKCKLVKFSWALVSATMAVTTVLTENTALAR 43
RESULT 5
AAR88823
ID AAR88823 standard; Protein; 271 AA.
XX AC AAR88823;
XX DT 25-JUN-1996 (first entry)
XX DE S. pyogenes DNaseB and leader sequence.
XX KM DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;
XX diagnosis; Escherichia coli.
XX OS Streptococcus pyogenes strain ATCC 14289.
XX FH Key Location/Qualifiers
FT Peptide 1..43
FT Protein /label= Sig_peptide
FT /label= 44..271
FT /label= Mat_protein
XX PN WO9606174-A1.
```

```
XX XX 29-FEB-1996.
XX PD 18-AUG-1994; 94WO-US09450.
XX PF 18-AUG-1994; 94WO-US09450.
XX PR 18-AUG-1994; 94WO-US09450.
XX PA (BECKI ) BECKMAN INSTR INC.
XX PI Adams CW, Belei MC, Pang PPY;
XX DR WPI; 1996-151377/15.
XX DR N-PSDB; AAT12774.
XX CC New DNA encoding Streptococcus pyogenes DNase B - for recombinant
PT prodn. of the enzyme in other bacteria, useful in immunoassays or
PT for treating cystic fibrosis
XX PS Claim 1; Page 67-70; 115pp; English.
XX CC Streptococcus pyogenes DNase B, including the leader peptide,
CC has the amino acid sequence given in AAR88821. The enzyme can
CC be obt'd. on a large scale by expression of encoding DNA (AAT12774)
CC in transformed host cells, esp. Escherichia coli. Inclusion
CC of the leader peptide facilitates purification of the recombinant
CC enzyme. The DNase B is useful in immunoassays to detect
CC anti-DNase B antibodies in serum as a marker for S. pyogenes
CC infection, and is also useful as a vaccine or for treatment, via
CC aerosol delivery, of cystic fibrosis.
XX SQ Sequence 271 AA;
SQ Query Match 100.0%; Score 203; DB 17; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNLGSRVRFSSKCKLVKFSWALVSATMAVTTVLTENTALAR 43
Db 1 MNLGSRVRFSSKCKLVKFSWALVSATMAVTTVLTENTALAR 43
RESULT 6
ABP30017
ID ABP30017 standard; Protein; 271 AA.
XX AC ABP30017;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 9210.
XX KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KM group A streptococcus; Streptococcus pyogenes; antibacterial;
KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus pyogenes.
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB04789.
XX PR 27-OCT-2000; 2000GB-0026333.
XX PR 24-NOV-2000; 2000GB-0028727.
XX PR 07-MAR-2001; 2001GB-0005640.
XX PA (CHIR-) CHIRON SPA.
XX PA (GENO-) INST GENOMIC RES.
XX PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Frazer C;
XX Tetteijn H;
```

DR WPI, 2002-352536/38.  
 DR N-PSDB; ABBN70648.  
 XX  
 PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX  
 PS Claim 1, Page 4045; 45255p; English.  
 XX  
 CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5463 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and anti-inflammatory  
 CC activity. (I), nucleic acids encoding (I), ABBN6044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 CC  
 SQ Sequence 271 AA;  
 QY  
 Db 1 MNLGSRVFSKCKRLVFSWALVSATMAVTTVLENTALAR 43  
 1 MNLGSRVFSKCKRLVFSWALVSATMAVTTVLENTALAR 43  
 RESULT 7  
 AAR70702  
 ID AAR70702 standard; Protein; 293 AA.  
 XX  
 AC AAR70702;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 15-AUG-1995 (first entry)  
 XX  
 DE DNA-ase-B.  
 XX  
 KM DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.  
 XX  
 OS Streptococcus pyogenes (ATCC 14289).  
 XX  
 FH Key  
 FH Protein Location/Qualifiers  
 FT 46..293  
 FT /note="mature protein"  
 FT Peptide 1..45  
 FT /note="leader peptide: claim 11"  
 FT misc\_difference 274  
 FT /note="in-frame stop codon"  
 XX  
 PN MO9500650-A1.  
 XX  
 PD 05-JAN-1995.  
 XX  
 PF 18-MAY-1994; 94WO-US05626.  
 XX  
 PR 23-JUN-1993; 93US-0082845.  
 XX  
 PA (BECT) BECKMAN INSTR INC.  
 XX  
 PI Adams CW, Belet CM, Pang PP;  
 XX

DR WPI, 1995-052087/07.  
 DR N-PSDB; AAG85037.  
 XX  
 PT New DNA encoding Streptococcus pyogenes DNase B - for diagnosing  
 PT S. pyogenes infection, also new promoter for expressing other  
 PT proteins  
 XX  
 PS Disclosure; Fig 5; 97pp; English.  
 XX  
 CC The sequence shows a gene product corresponding to a Streptococcus  
 CC pyogenes DNA-ase-B. The protein is useful as a diagnostic agent,  
 CC vaccine or as an aerosol to treat excessive lung viscosity, e.g. in  
 CC cystic fibrosis.  
 CC (Updated on 25-MAR-2003 to correct PW field.)  
 CC  
 SQ Sequence 293 AA;  
 QY  
 Db 2 NLLGSRVFSKCKRLVFSWALVSATMAVTTVLENTALAR 43  
 4 NLLGSRVFSKCKRLVFSWALVSATMAVTTVLENTALAR 45  
 RESULT 8  
 AAU00294  
 ID AAU00294 standard; Protein; 1589 AA.  
 XX  
 AC AAU00294;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Interferon induced polypeptide, IFN2.  
 XX  
 KM Interferon induced nucleic acid; autoimmune disease; lupus erythematosus;  
 KM immunodeficiency; stroke; acquired immunodeficiency syndrome; AIDS;  
 KM graft rejection; viral infection; hepatitis; aplastic anaemia; cancer;  
 KM human immunodeficiency virus; HIV; immune-mediated glomerulonephritis;  
 KM haematologic disease; chronic neutropenia; myocardial infarction;  
 KM neurological disease; Alzheimer's disease; Parkinson's disease;  
 KM amyotrophic lateral sclerosis; spinal muscular atrophy; human; IFN2.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WC200118208-A2.  
 XX  
 PD 15-MAR-2001.  
 XX  
 PF 08-SEP-2000; 2000WO-US24704.  
 XX  
 PR 08-SEP-1999; 99US-0152921.  
 PR 20-OCT-1999; 99US-0160575.  
 PR 20-JAN-2000; 2000US-0177104.  
 PR 07-SEP-2000; 2000US-0656633.  
 XX  
 PA (CURA-) CUPAGEN CORP.  
 PA (BIO) BIOGEN INC.  
 XX  
 PI Peyman JA, Da Silva A, Hochman P, Hsu A;  
 XX  
 DR WPI, 2001-235201/24.  
 DR N-PSDB; AAS01147.  
 XX  
 PT New interferon induced polypeptides and polynucleotides, useful for the  
 PT diagnosis, prevention and treatment of immunological, cell  
 PT proliferative disorders, such as lupus erythematosus, cancer, stroke  
 PT and Alzheimer's disease -  
 XX  
 PS Claim 1, Page 17-22; 134pp; English.  
 CC  
 CC The sequence represents interferon induced polypeptide, IFN2. IFN

CC nucleic acids and polypeptides are useful for treating or preventing a  
 CC pathology associated with IFN polypeptide in a human. They are useful for  
 CC determining the presence of or predisposition to a disease associated  
 CC with altered levels of IFN polypeptide or polynucleotide. IFN nucleic  
 CC acids, polypeptides and antibodies are useful for diagnosis, prevention  
 CC or treatment of variety of immunological and cell proliferative  
 CC disorders, such as autoimmune diseases e.g. lupus erythematosus,  
 CC immunodeficiency diseases such as acquired immunodeficiency syndrome  
 CC (AIDS), graft rejection, viral infections including hepatitis and human  
 CC immunodeficiency virus (HIV), immune-mediated glomerulonephritis,  
 CC haematologic diseases such as aplastic anaemia and chronic neutropenia  
 CC and cancer. In addition they are also useful for treating or diagnosing  
 CC various disorders associated with cell death, including myocardial  
 CC infarction, stroke, neurological diseases including Alzheimer's and  
 CC Parkinson's diseases, amyotrophic lateral sclerosis and spinal muscular  
 CC atrophy. IFN nucleic acids and polypeptides are also useful for  
 CC identifying interferon-like proteins and interferon agonists, for  
 CC screening drugs and compounds which inhibit or enhance IFN activity or  
 CC function and as targets for the identification of small molecules that  
 CC are immunostimulatory, immunosuppressive, or stimulate or suppress normal  
 CC cell or tumour cell growth in mammals, including humans.

SO Sequence 1589 AA;  
 Query Match 28.1%; Score 57; DB 22; Length 1589;  
 Best Local Similarity 34.3%; Pred. No. 21;  
 Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;  
 QY 2 NLLGSRVRFSSKCRIVKFSMVALVSATMAVTTVTL 36  
 DB 903 NILKQNIPTKAKL--FSFLALNSYVDDTISL 935

RESULT 9  
 ABG52177  
 ID ABG52177 standard; Peptide; 1591 AA.

AC ABG52177;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 30825.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.

OS Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488898/53.

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human adult liver -  
 XX Claim 27; SEQ ID No 30825; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (1) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 1109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult  
 CC liver. (1) may be used for predicting, measuring and displaying gene  
 CC expression in samples derived from human adult liver. The genes  
 CC identified may be involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which  
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent  
 CC human liver single exon encoded peptides of the invention.  
 CC Note: The sequence information for this patent does not appear in the  
 CC printed specification but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 1591 AA;  
 Query Match 28.1%; Score 57; DB 22; Length 1591;  
 Best Local Similarity 34.3%; Pred. No. 21;  
 Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;  
 QY 2 NLLGSRVRFSSKCRIVKFSMVALVSATMAVTTVTL 36  
 DB 905 NILKQNIPTKAKL--FSFLALNSYVDDTISL 937

RESULT 10  
 ID ABB32113 standard; Peptide; 1591 AA.

AC ABB32113;

XX 01-FEB-2002 (first entry)

XX Peptide #4764 encoded by breast cell single exon nucleic acid probe.

KW Human; microarray; single exon probe; gene expression; breast;  
 KW disease; cancer.

OS Homo sapiens.

XX WO200157271-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00662.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-496933/54.

PT New spatially-addressable set of single exon nucleic acid probes,  
 PT useful for measuring gene expression in sample derived from human  
 PT breast, comprises number of single exon nucleic acid probes -  
 XX Claim 27; SEQ ID NO 15081; 327bp + sequence listing; English.

CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting  
 CC the probes with a collection of detectably labelled nucleic acids  
 CC derived from mRNA of human breast, and then measuring the label  
 CC bound to each probe of the microarray. The probes are useful for

CC	verifying the expression of regions of genomic DNA predicted to
CC	encode proteins. They are useful for gene discovery, and for
CC	determining predisposition and/or prognosing breast disease. Gene
CC	expression analysis is useful for assessing the toxicity of chemical
CC	agents on cells. The microarray of this invention presents a far greater
CC	diversity of probes for measuring gene expression, with far less bias
CC	than expressed sequence tag microarrays. The method is suitable for
CC	rapid production of functional information from genomic sequence. The
CC	present sequence is a peptide encoded by a single exon nucleic acid
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 1591 AA;
QY	2 NLLGSRRVFSKCCRLVKFSGMALVSAATMAVTVTL 36
Db	905 NILKGNITFKKAL--FSFLALINSYVDTTISL 937
XX	
RESULT 11	
AB37365	
ID	AB37365 standard; Peptide; 1591 AA.
XX	
AC	AB37365;
XX	
DT	04-FEB-2002 (first entry)
XX	
DE	Peptide #4871 encoded by human foetal liver single exon probe.
XX	
KW	Human; foetal liver; gene expression; single exon nucleic acid probe.
XX	
OS	Homo sapiens.
XX	
PN	WO200157277-A2.
XX	
PD	09-AUG-2001.
XX	
PF	30-JAN-2001; 2001WO-US00669.
XX	
PR	04-FEB-2000; 2000US-0180312.
XX	
PR	26-MAY-2000; 2000US-0207456.
XX	
PR	30-JUN-2000; 2000US-0608408.
XX	
PR	03-AUG-2000; 2000US-0632366.
XX	
PR	21-SEP-2000; 2000US-0234687.
XX	
PR	27-SEP-2000; 2000US-0236359.
XX	
PR	04-OCT-2000; 2000GB-0024263.
XX	
PA	(MOL-)
XX	MOLECULAR DYNAMICS INC.
XX	
P1	Penn SG, Hanzel DK, Chen W, Rank DR;
XX	
DR	WPI; 2001-483447/52.
XX	
PT	Human genome-derived single exon nucleic acid probes useful for
XX	analyzing gene expression in human fetal liver -
XX	
PS	Claim 27; SEQ ID NO 30000; 639bp + sequence listing; English.
XX	
CC	The invention relates to a single exon nucleic acid probe for
CC	measuring human gene expression in a sample derived from human foetal
CC	liver. The single exon nucleic acid probes may be used for predicting,
CC	measuring and displaying gene expression in samples derived from human
CC	fetal liver. The present sequence is a peptide encoded by a single exon
CC	nucleic acid probe of the invention.
CC	Note: The sequence data for this patent did not form part of the
CC	printed specification, but was obtained in electronic format directly
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.

[illegible]

Db 905 NILKGNIPTKEAKL--FSFLALNSYVPTTISL 937

RESULT 13  
AAM58054  
ID AAM58054 standard; Protein, 1591 AA.

AC AAM58054;  
DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30159.

XX Human; brain expressed exon; gene expression analysis; probe;  
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00667.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

XX Example 4; SEQ ID NO: 30159; 650pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.

XX Sequence 1591 AA;

XX Query Match 28.1%; Score 57; DB 22; Length 1591;

XX Best Local Similarity 34.3%; Pred. No. 21;

XX Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;

XX 2 NILGSRVFSKCKRLVFSWALVSATMAVTYTL 36

XX Db 905 NILKGNIPTKEAKL--FSFLALNSYVPTTISL 937

RESULT 14

ID AAM70495 standard; Protein, 1591 AA.

AC AAM70495;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 30801.

XX Human; bone marrow expressed exon; gene expression analysis; probe;  
KW microarray; cancer; leukaemia; lymphoma; myeloma.

OS Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 30801; 658pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 1591 AA;

XX Query Match 28.1%; Score 57; DB 22; Length 1591;

XX Best Local Similarity 34.3%; Pred. No. 21;

XX Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;

XX 2 NILGSRVFSKCKRLVFSWALVSATMAVTYTL 36

XX Db 905 NILKGNIPTKEAKL--FSFLALNSYVPTTISL 937

RESULT 15

ID AAM18331 standard; Protein, 1591 AA.

AC AAM18331;

DT 12-OCT-2001 (first entry)

XX Peptide #4765 encoded by probe for measuring cervical gene expression.

XX Probe; human; microarray; gene expression; cervical epithelial cell;

XX cervical cancer.

OS Homo sapiens.

PN WO200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for  
 analyzing gene expression in human cervical epithelial cells -

PS Claim 27; SEQ ID No 23157; 487bp; English.

CC The present invention relates to human single exon nucleic acid probes  
 (SENP: see A110068-A128459). The present sequence is a peptide encoded  
 by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 can be used to produce a single exon microarray, which can be used for  
 measuring human gene expression in a sample derived from human cervical  
 epithelial cells. By measuring gene expression, the probes are therefore  
 useful in grading and/or staging of diseases of the cervix, notably  
 cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 specification, but was obtained in electronic format directly from WIPO  
 at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1591 AA;

Query Match 28.1%; Score 57; DB 22; Length 1591;

Best Local Similarity 34.3%; Pred No. 21;  
 Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;

QY 2 NLLGSRVFSKCKRLVKSMAVLSATMAVTTVL 36

DB 905 NLLKGONFTKEAKL--FSFLALNSYVPDITSL 937

Search completed: January 5, 2004, 18:38:05  
 Job time : 13.0866 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:35:50 ; Search time 4.90976 Seconds  
(without alignments)  
370.561 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203  
Sequence: 1 MALLGSRVPSKCRIVKFS.....LVSATMAVTLENTALAR 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	100.0	43	4	US-08-393-889-1
2	203	100.0	43	5	PCT-US94-09450-1
3	203	100.0	271	4	US-08-393-889-8
4	203	100.0	271	5	PCT-US94-09450-8
5	194	95.6	272	4	US-08-393-889-15
6	194	95.6	272	5	PCT-US94-09450-15
7	55.5	27.3	397	1	US-08-290-448A-76
8	55.5	27.3	397	1	US-08-290-448A-76
9	55.5	27.3	397	1	US-08-175-069A-76
10	55.5	27.3	397	4	US-08-461-939B-76
11	55.5	27.3	397	4	US-08-464-000-76
12	54	26.6	927	4	US-09-198-452A-472
13	50	24.6	50	2	US-08-467-046-29
14	50	24.6	604	4	US-07-690-983D-12
15	50	24.6	604	4	US-09-519-232-64
16	49.5	24.4	387	1	US-08-290-448A-72
17	49.5	24.4	387	1	US-08-290-448A-72
18	49.5	24.4	387	1	US-08-175-069A-72
19	49.5	24.4	387	4	US-08-461-939B-72
20	49.5	24.4	387	4	US-08-464-000-72
21	49.5	24.4	394	3	US-09-390-721-5
22	49.5	24.4	571	2	US-08-796-414B-1
23	49	24.1	600	3	US-09-212-971-12
24	49	24.1	600	3	US-08-800-929A-12
25	49	24.1	600	4	US-09-617-053A-12
26	49	24.1	602	4	US-09-201-936-40
27	48	23.6	105	4	US-09-732-210-130

28	48	23.6	241	1	US-07-987-286-17	Sequence 17, Appl
29	48	23.6	241	2	US-08-614-626-17	Sequence 17, Appl
30	46	22.7	181	4	US-09-252-991A-25104	Sequence 25104, A
31	46	22.7	492	4	US-09-134-001C-4847	Sequence 4847, Ap
32	46	22.7	533	4	US-09-252-991A-22953	Sequence 22953, A
33	46	22.7	602	4	US-09-134-001C-4416	Sequence 4416, Ap
34	46	22.7	1391	4	US-09-106-568E-8	Sequence 8, Appl
35	45.5	22.4	253	1	US-08-265-087-4	Sequence 4, Appl
36	45.5	22.4	253	1	US-08-621-493-4	Sequence 4, Appl
37	45.5	22.4	253	2	US-08-684-687-4	Sequence 4, Appl
38	45.5	22.4	253	2	US-08-965-688-4	Sequence 4, Appl
39	45.5	22.4	253	3	US-09-260-173-6	Sequence 4, Appl
40	45.5	22.4	253	4	US-09-924-703-6	Sequence 6, Appl
41	45	22.2	430	4	US-09-252-991A-30239	Sequence 1, Appl
42	45	22.2	513	1	US-08-403-866-1	Sequence 10239, A
43	45	22.2	568	4	US-09-252-991A-17693	Sequence 17693, A
44	45	22.2	932	4	US-09-252-991A-26175	Sequence 26175, A
45	44.5	21.9	228	4	US-09-198-452A-1168	Sequence 1168, Ap

#### ALIGNMENTS

RESULT 1  
US-08-393-889-1  
; Sequence 1, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Bellet, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESS: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393, 889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082, 845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 795-6321  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
US-08-393-889-1  
Query Match 100.0%; Score 203; DB 4; Length 43;

Best Local Similarity 100.0%; Pred. No. 1.2e-23;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43

## RESULT 2

PCT-US94-09450-1  
Sequence 1, Application PC/TUS9409450

GENERAL INFORMATION:  
APPLICANT: Beckman Instruments, Inc.  
APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant Dnase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
PCT-US94-09450-1

Query Match 100.0%; Score 203; DB 5; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.2e-23;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43

## RESULT 3

US-08-393-889-8  
Sequence 8, Application US/08393889

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant Dnase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-393-889-8

Query Match 100.0%; Score 203; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTLENTALAR 43

## RESULT 4

PCT-US94-09450-8  
Sequence 8, Application PC/TUS9409450

GENERAL INFORMATION:  
APPLICANT: Beckman Instruments, Inc.  
APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant Dnase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-09450-8

Query Match 100.0%; Score 203; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.1e-22;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 43  
Db 1 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 43

## RESULT 5

US-08-393-889-15  
Sequence 15, Application US/08393889  
Patent No. 6420152

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belal, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
NUMBER OF SEQUENCES: 16  
TITLE OF INVENTION: Streptococcus pyogenes  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fairber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-393-889-15

Query Match 95.6%; Score 194; DB 4; Length 272;  
Best Local Similarity 97.6%; Pred. No. 2.5e-21;

Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 43  
Db 4 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 45

## RESULT 6

PCT-US94-09450-15  
Sequence 15, Application PC/TUS9409450  
GENERAL INFORMATION:

APPLICANT: Beckman Instruments, Inc.  
APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-09450-15

Query Match 95.6%; Score 194; DB 5; Length 272;  
Best Local Similarity 97.6%; Pred. No. 2.5e-21;  
Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 43  
Db 4 NLLGSRVFSKCKRLVKFSWALVSATMAVTTVLTENTALAR 45

## RESULT 7

US-08-290-448A-76  
Sequence 76, Application US/08290448A  
Patent No. 5676954

GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,448A  
FILING DATE: August 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-290-448A-76

Query Match 27.3%; Score 55.5; DB 1; Length 397;  
Best Local Similarity 36.8%; Pred. No. 2.4; Mismatches 16; Indels 1; Gaps 1;  
Matches 14; Conservative 7;

QY 1 MNLGSRVFSKKCRIVKFSMVALVSATNAVTTVTLEN 38  
DB 198 INVAGSSQIWDHCSLSK-SFDGLVDVTLGSTHTVTISN 234

RESULT 8  
US-08-290-448A-76  
Sequence 76, Application US/08290448A  
Patent No. 5698204  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/290,448A  
FILING DATE: August 15, 1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-290-448A-76

Query Match 27.3%; Score 55.5; DB 1; Length 397;  
Best Local Similarity 36.8%; Pred. No. 2.4; Mismatches 16; Indels 1; Gaps 1;  
Matches 14; Conservative 7;

QY 1 MNLGSRVFSKKCRIVKFSMVALVSATNAVTTVTLEN 38  
DB 198 INVAGSSQIWDHCSLSK-SFDGLVDVTLGSTHTVTISN 234

RESULT 9  
US-08-175-069A-76  
Sequence 76, Application US/08175069A  
Patent No. 5776761  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Rafnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/175,069A  
FILING DATE: December 29, 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: May 29, 1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: March 17, 1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018DV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-175-069A-76

Query Match 27.3%; Score 55.5; DB 1; Length 397;  
Best Local Similarity 36.8%; Pred. No. 2.4; Mismatches 16; Indels 1; Gaps 1;  
Matches 14; Conservative 7;

QY 1 MNLGSRVPSKRCRLVFSMVALVSATMAVTTTLEN 38  
Db 198 INVAGSSQIWMIDHCSLSK-SFDGLVDVTLGSTHTVTSN 234

RESULT 10  
US-08-461-939B-76  
Sequence 76, Application US/08461939B  
Patent No. 6335019  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.  
APPLICANT: Ratnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Methods For Treating Sensitivity To A  
TITLE OF INVENTION: Protein Allergen Using Peptides Which Include A T Cell Epitope  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 28 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,939B  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/464,000  
FILING DATE: 05-JUN-1995  
APPLICATION NUMBER: US 08/290,448  
FILING DATE: 15-AUG-1994  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: 29-MAY-1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: 17-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CNDV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)742-4214  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-939B-76

Query Match 27.3%; Score 55.5; DB 4; Length 397;  
Best Local Similarity 36.8%; Pred. No. 2.4;  
Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNLGSRVPSKRCRLVFSMVALVSATMAVTTTLEN 38  
Db 198 INVAGSSQIWMIDHCSLSK-SFDGLVDVTLGSTHTVTSN 234

RESULT 11  
US-08-464-000-76  
Sequence 76, Application US/08464000  
Patent No. 6335020  
GENERAL INFORMATION:  
APPLICANT: Rogers, Bruce  
APPLICANT: Klapper, David G.

APPLICANT: Ratnar, Thorunn  
APPLICANT: Kuo, Mei-chang  
TITLE OF INVENTION: Allergenic Peptides from Ragweed Pollen  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD, LLP  
STREET: 60 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,000  
FILING DATE: 05-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/290,448  
FILING DATE: 15-AUG-1994  
APPLICATION NUMBER: US 07/529,951  
FILING DATE: 29-MAY-1990  
APPLICATION NUMBER: US 07/325,365  
FILING DATE: 17-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Amy E. Mandragouras  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: IMI-018CND2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 76:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-000-76

Query Match 27.3%; Score 55.5; DB 4; Length 397;  
Best Local Similarity 36.8%; Pred. No. 2.4;  
Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNLGSRVPSKRCRLVFSMVALVSATMAVTTTLEN 38  
Db 198 INVAGSSQIWMIDHCSLSK-SFDGLVDVTLGSTHTVTSN 234

RESULT 12  
US-09-198-452A-472  
Sequence 472, Application US/09198452A  
Patent No. 6559294  
GENERAL INFORMATION:  
APPLICANT: Griffiths, R.  
TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments  
TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention  
FILE REFERENCE: 9710-003-999  
CURRENT APPLICATION NUMBER: US/09/198,452A  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 6849  
SEQ ID NO 472  
LENGTH: 927  
TYPE: PRT  
ORGANISM: Chlamydia pneumoniae  
FEATURE:  
NAME/KEY: SITE  
LOCATION: 1...927  
OTHER INFORMATION: Xaa=unknown or other  
US-09-198-452A-472

Query Match 26.6%; Score 54; DB 4; Length 927;  
Best Local Similarity 34.3%; Pred. No. 11;  
Matches 12; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

OY 5 GSRVFSKCKRLVKSMTALVSATMAVTTTLENT 39  
DB 483 GSLSTICGCRFLKRDQEXTLKAQASQTTLTSSG 517

RESULT 13  
US-08-467-046-29  
; Sequence 29, Application US/08467046  
; Patent No. 5948644  
; GENERAL INFORMATION:  
; APPLICANT: DOPHEIDE, THEODOREUS AA  
; APPLICANT: FRENKEL, MARICE J  
; APPLICANT: GRANT, WARWICK N  
; APPLICANT: SAVIN, KEITH W  
; APPLICANT: WAGLAND, BARRY M  
; TITLE OF INVENTION: VACCINE  
; NUMBER OF SEQUENCES: 32  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: SUITE 500, 1800 DIAGONAL ROAD  
; CITY: ALEXANDRIA  
; STATE: VIRGINIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 22113-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,046  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/904,055  
; FILING DATE: 26-JUN-1992  
; APPLICATION NUMBER: AU P00621  
; FILING DATE: 26-SEP-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU P00622  
; FILING DATE: 26-SEP-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU P00623  
; FILING DATE: 26-SEP-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU P00624  
; FILING DATE: 26-SEP-1988  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU89/00416  
; FILING DATE: 26-SEP-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/548,901  
; FILING DATE: 26-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, STEPHEN A  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16786/166  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-9300  
; TELEFAX: (703) 836-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 50 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-467-046-29

Query Match 24.6%; Score 50; DB 2; Length 50;  
Best Local Similarity 37.9%; Pred. No. 1.4;  
Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

OY 15 RLVKFSMTALVSATMAVTTTLENTALAR 43  
DB 2 KMKKLMVVALVSTLALSGGAMSTAIKK 30

RESULT 14  
US-07-690-983D-12  
; Sequence 12, Application US/07690983D  
; Patent No. 5403586  
; GENERAL INFORMATION:  
; APPLICANT: RUSSELL-JONES, Gregory J.  
; APPLICANT: STEWART, Andrew G.  
; APPLICANT: TSONIS, Con G.  
; TITLE OF INVENTION: FUSION PROTEINS  
; NUMBER OF SEQUENCES: 47  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/690,983D  
; FILING DATE: 25-JUN-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/AU90/00373  
; FILING DATE: 24-AUG-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 16786/148 CHAC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 243 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-690-983D-12

Query Match 24.6%; Score 50; DB 1; Length 243;  
Best Local Similarity 37.9%; Pred. No. 9;  
Matches 11; Conservative 9; Mismatches 9; Indels 0; Gaps 0;

OY 15 RLVKFSMTALVSATMAVTTTLENTALAR 43  
DB 2 KMKKLMVVALVSTLALSGGAMSTAIKK 30

RESULT 15  
US-09-519-232-64  
; Sequence 64, Application US/09519232  
; Patent No. 6528702  
; GENERAL INFORMATION:  
; APPLICANT: Salmeron, John  
; APPLICANT: Weisio, Laura  
; APPLICANT: Willits, Michael  
; APPLICANT: Mengiste, Tesfaye  
; TITLE OF INVENTION: NOVEL PLANT GENES AND USES THEREOF  
; FILE REFERENCE: S-30857A/RTP2095  
; CURRENT APPLICATION NUMBER: US/09/519,232



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 5, 2004, 18:41:44 ; Search time 79.6648 seconds  
(without alignments)  
108.298 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203  
Sequence: 1 MNLGSRKRVFSKCRIVKFS.....IVSATMAYTVLTLENTALAR 43

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203	100.0	43	9	US-09-119-900-1
2	203	100.0	271	9	US-09-119-900-8
3	194	95.6	272	9	US-09-119-900-15
4	57	28.1	1591	9	US-09-864-761-37952
5	55.5	27.3	397	11	US-09-847-208-15
6	52.5	25.9	1019	12	US-10-369-493-5042
7	51	25.1	481	9	US-09-767-041-12
8	50	24.6	604	12	US-10-328-675A-64
9	49.5	24.4	396	11	US-09-847-208-13
10	49.5	24.4	728	12	US-10-287-274-300
11	49	24.1	75	15	US-10-129-851-5
12	49	24.1	178	8	US-08-781-986A-5214
13	49	24.1	282	12	US-10-104-047-3069
14	49	24.1	299	11	US-09-832-522-74
15	49	24.1	303	11	US-09-832-522-60

16	49	24.1	321	12	US-09-912-976-67	Sequence 67, Appl
17	49	24.1	321	12	US-10-005-041A-71	Sequence 71, Appl
18	49	24.1	321	15	US-10-219-834-72	Sequence 72, Appl
19	49	24.1	600	10	US-09-974-592-12	Sequence 12, Appl
20	49	24.1	602	10	US-09-201-936-40	Sequence 14, Appl
21	48.5	23.9	363	15	US-10-156-761-13425	Sequence 13425, A
22	48.5	23.9	467	12	US-10-032-585-7233	Sequence 7233, Ap
23	48	23.6	252	10	US-09-738-626-4091	Sequence 4091, Ap
24	48	23.6	420	12	US-10-369-493-8086	Sequence 8086, Ap
25	48	23.6	450	12	US-10-369-493-2191	Sequence 2191, Ap
26	48	23.6	661	12	US-10-281-868-2	Sequence 2, Appl
27	48	23.6	661	12	US-10-391-399-17	Sequence 17, Appl
28	48	23.6	661	14	US-10-094-214-5	Sequence 5, Appl
29	47	23.2	47	10	US-09-738-626-6139	Sequence 6139, Ap
30	47	23.2	65	15	US-10-106-698-7381	Sequence 7381, Ap
31	47	23.2	494	9	US-09-741-669-468	Sequence 468, App
32	47	23.2	849	12	US-10-369-493-4999	Sequence 4999, Ap
33	46	22.7	204	16	US-10-080-170-561	Sequence 561, App
34	46	22.7	424	9	US-09-815-242-11177	Sequence 11177, A
35	46	22.7	430	12	US-10-369-493-12210	Sequence 12210, A
36	46	22.7	1391	11	US-09-994-595-8	Sequence 8, Appl
37	45.5	22.4	253	10	US-09-924-703-6	Sequence 6, Appl
38	45.5	22.4	297	15	US-10-103-197-5	Sequence 5, Appl
39	45.5	22.4	547	12	US-10-032-585-7584	Sequence 7584, Ap
40	45	22.2	491	10	US-09-738-626-4423	Sequence 4423, Ap
41	45	22.2	513	12	US-10-306-905-13	Sequence 13, Appl
42	45	22.2	563	12	US-10-369-493-12461	Sequence 12461, A
43	45	22.2	577	12	US-10-369-493-10270	Sequence 10270, A
44	45	22.2	819	12	US-10-238-075-933	Sequence 933, App
45	45	22.2	844	12	US-10-369-493-1193	Sequence 1193, Ap

#### ALIGNMENTS

RESULT 1  
US-09-119-900-1  
; Sequence 1, Application US/09119900  
; Patent No. US0020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-1

Query Match 100.0%; Score 203; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 3.8e-22;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43  
DB 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43

## RESULT 2

US-09-119-900-8  
Sequence 8, Application US/09119900  
Patent No. US20020081622A1

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belej, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-119-900-8

Query Match 100.0%; Score 203; DB 9; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.2e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43  
DB 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43

## RESULT 3

US-09-119-900-15  
Sequence 15, Application US/09119900  
Patent No. US20020081622A1

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belej, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-119-900-15

Query Match 95.6%; Score 194; DB 9; Length 272;  
Best Local Similarity 97.6%; Pred. No. 6.5e-20;  
Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43  
DB 4 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 45

## RESULT 4

US-09-864-761-37952  
Sequence 37952, Application US/09864761  
Patent No. US20020048763A1

GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26

QY 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43  
DB 1 NMLGSRVFSKCKRLVKFSWALVSATMAVTTTLENTALAR 43



```

; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annamex Sequence Listing Engine vers. 1.1
; SEQ ID NO 37952
; LENGTH: 1591
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000119.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 7.7
; OTHER INFORMATION: EST HUMAN HIT: BE888763.1, EVALUATE 1.00e-113
; OTHER INFORMATION: SWISSPROT HIT: P50495, EVALUATE 1.80e-02
; US-09-864-761-37952

Query Match      28.1%; Score 57; DB 9; Length 1591;
Best Local Similarity 34.3%; Pred. No. 31;
Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;

Cy      2 NLLGSRVFSKRCRLVKSFWALVSATMAVTVTTL 36
Db      905 NILKQNIPTKAKL--FSPFLALNSVDPDTTSL 937

RESULT 5
US-09-847-208-15
; Sequence 15, Application US/09847208
; Publication No. US20030082190A1
; GENERAL INFORMATION:
; APPLICANT: Saxon, Andrew
; APPLICANT: Zhang, Ke
; APPLICANT: Zhu, Daocheng
; TITLE OF INVENTION: FUSE-MEDIATED ALLERGIC DISEASES
; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
```

```

; FILE REFERENCE: UC67.002A
; CURRENT APPLICATION NUMBER: US/09/847,208
; CURRENT FILING DATE: 2001-05-01
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 15
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Ambrosia artemisiifolia (Short ragweed)
; US-09-847-208-15

Query Match      27.3%; Score 55.5; DB 11; Length 397;
Best Local Similarity 36.8%; Pred. No. 10;
Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

Cy      1 NMLGSRVFSKRCRLVKSFWALVSATMAVTVTTL 38
Db      198 INVAGSSQIWDHCSLSK-SFDGLVDVTLGSTHVITSN 234

RESULT 6
US-10-369-493-5042
; Sequence 5042, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Cao, Yongwei
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5042
; LENGTH: 1019
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; US-10-369-493-5042

Query Match      25.9%; Score 52.5; DB 12; Length 1019;
Best Local Similarity 29.5%; Pred. No. 82;
Matches 13; Conservative 11; Mismatches 5; Indels 15; Gaps 2;

Cy      8 RVFSKCR-----LVKFSWVA-----LVSATMAVTVTTL 36
Db      575 RFRRCRHPGSGLSQILKFTPLSGARKVIISTNIAETSVTI 618

RESULT 7
US-09-767-041-12
; Sequence 12, Application US/09767041
; Patent No. US20020055168A1
; GENERAL INFORMATION:
; APPLICANT: Smith, Hilda
; TITLE OF INVENTION: STEPTOCOCCUS SUIIS VACCINES AND DIAGNOSTIC TESTS
; FILE REFERENCE: 2183-4726
; CURRENT APPLICATION NUMBER: US/09/767,041
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: PCT/NL99/00460
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: EP98202465.5
; PRIOR FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: EP98202467.1
; PRIOR FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 481
```

TYPE: PRT  
ORGANISM: Streptococcus suis  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: CPS2A  
US-09-767-041-12

Query Match 25.1%; Score 51; DB 9; Length 481;  
Best Local Similarity 38.7%; Pred. No. 57;  
Matches 12; Conservative 7; Mismatches 12; Indels 0; Gaps 0;

QY 6 SRRVFSKRCRLVFSMVALVSATMAVTTVTL 36  
DB 5 SGRSKSKKFLVNFALGYSITLCLFLVTM 35

RESULT 8  
US-10-328-675A-64  
Sequence 64, Application US/10328675A  
Publication No. US20030159171A1  
GENERAL INFORMATION:  
APPLICANT: Salmeron, John  
APPLICANT: Weislo, Laura  
APPLICANT: Willets, Michael  
TITLE OF INVENTION: NOVEL PLANT GENES AND USES THEREOF  
FILE REFERENCE: 30857USNPDI1  
CURRENT APPLICATION NUMBER: US/10/328,675A  
CURRENT FILING DATE: 2002-12-23  
PRIOR APPLICATION NUMBER: 09/519,232  
PRIOR FILING DATE: 2000-03-06  
PRIOR APPLICATION NUMBER: 60/219,338  
PRIOR FILING DATE: 1999-03-09  
NUMBER OF SEQ ID NOS: 74  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 64  
LENGTH: 604  
TYPE: PRT  
ORGANISM: Beta vulgaris  
US-10-328-675A-64

Query Match 24.6%; Score 50; DB 12; Length 604;  
Best Local Similarity 32.4%; Pred. No. 1e+02;  
Matches 12; Conservative 6; Mismatches 19; Indels 0; Gaps 0;

QY 3 LIGSRVFSKRCRLVFSMVALVSATMAVTTVTL 39  
DB 9 VIDSRTAFSDSNDISNGSSICCVATTTTAAENS 45

RESULT 9  
US-09-847-208-13  
Sequence 13, Application US/09847208  
Publication No. US20030082190A1  
GENERAL INFORMATION:  
APPLICANT: Saxon, Andrew  
APPLICANT: Zhang, Ke  
APPLICANT: Zhu, Daoceng  
TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF  
FILE REFERENCE: UCG7.002A  
CURRENT APPLICATION NUMBER: US/09/847,208  
CURRENT FILING DATE: 2001-05-01  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Ambrosia artemisiifolia (Short ragweed)  
US-09-847-208-13

Query Match 24.4%; Score 49.5; DB 11; Length 396;  
Best Local Similarity 30.8%; Pred. No. 74;  
Matches 12; Conservative 11; Mismatches 15; Indels 1; Gaps 1;

QY 1 MMLGSRVFSKRCRLVFSMVALVSATMAVTTVTL 39  
DB 197 ISIGSSQIWDHCLSK-SVDGLVADLGTTRLVNS 234

RESULT 10  
US-10-287-274-300  
Sequence 300, Application US/10287274  
Publication No. US20030181408A1  
GENERAL INFORMATION:  
APPLICANT: Foreyth, R. Allen  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zykaid, Judith  
TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERETO  
FILE REFERENCE: ELITRA.008DV1  
CURRENT APPLICATION NUMBER: US/10/287,274  
CURRENT FILING DATE: 2002-10-31  
PRIOR APPLICATION NUMBER: US-60/164415  
PRIOR FILING DATE: 1999-11-09  
PRIOR APPLICATION NUMBER: US 09/711164  
PRIOR FILING DATE: 2000-11-09  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 300  
LENGTH: 728  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-10-287-274-300

Query Match 24.4%; Score 49.5; DB 12; Length 728;  
Best Local Similarity 34.2%; Pred. No. 1.5e+02;  
Matches 13; Conservative 7; Mismatches 17; Indels 1; Gaps 1;

QY 7 RRVFSKRCRLVFSMVALVSATMAVTTVTL-ENTALAR 43  
DB 60 RHFSNKDFVKFSTSVLQADKEITLTKNNENTTLTO 97

RESULT 11  
US-10-129-851-5  
Sequence 5, Application US/10129851  
Publication No. US20030120039A1  
GENERAL INFORMATION:  
APPLICANT: Xuan Mei Li  
APPLICANT: Paul Robert Roestek, Jr.  
APPLICANT: Eric Wen Su  
APPLICANT: He Wang  
APPLICANT: Jian-Jun Wei  
TITLE OF INVENTION: Human Preoptic Regulatory Factor-2 and Uses Thereof  
FILE REFERENCE: X-12919  
CURRENT APPLICATION NUMBER: US/10/129,851  
CURRENT FILING DATE: 2002-05-08  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 5  
LENGTH: 75  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-129-851-5

Query Match 24.1%; Score 49; DB 15; Length 75;  
Best Local Similarity 35.7%; Pred. No. 13;  
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

QY 14 CRLVFSMVALVSATMAVTTVTLTENTAL 41  
DB 4 CYLIRFLQVFPVOPANVAVTMDVSNLNM 31

RESULT 12  
US-08-781-986A-5214  
Sequence 5214, Application US/08781986A

```
Publication No. US20030054436A1
GENERAL INFORMATION:
APPLICANT: Charles Kunech
TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
NUMBER OF SEQUENCES: 5255
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 5214:
SEQUENCE CHARACTERISTICS:
LENGTH: 178 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-781-986A-5214

Query Match 24.1%; Score 49; DB 8; Length 178;
Best Local Similarity 33.3%; Pred. No. 35;
Matches 11; Conservative 10; Mismatches 10; Indels 2; Gaps 1;

Cy 8 RVFSKCKLVKFSWALVSATMAVTVTLENTA 40
Db 10 RIINMKNLIAKSLTL--AAIGITTTTASTA 40

RESULT 13
US-10-104-047-3069
Sequence 3069, Application US/10104047
Publication No. US20030236392A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20030236392A1 full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3069
LENGTH: 282
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-3069

Query Match 24.1%; Score 49; DB 12; Length 282;
Best Local Similarity 35.7%; Pred. No. 59;
Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;
```

```
Cy 14 CRLVFSWALVSATMAVTVTLENTAL 41
Db 211 CYLIRFLQVFPQANVAATKMDVSNLAM 238

RESULT 14
US-09-832-522-74
Sequence 74, Application US/09832522
Publication No. US20030091563A1
GENERAL INFORMATION:
APPLICANT: Shenoy, Suresh G
APPLICANT: Gangoli, Esna A
APPLICANT: Raestelli, Luca
APPLICANT: Smithson, Glenda
APPLICANT: Padigara, Muralidhara
APPLICANT: Verne, Corine
APPLICANT: Wolenc, Adam R
APPLICANT: Caeman, Stacie J
APPLICANT: Tchernen, Velizar T
APPLICANT: Szekeres, Edward S
APPLICANT: Grose, William
APPLICANT: Alsebrook, John P
APPLICANT: Burgess, Catherine E
TITLE OF INVENTION: No. US20030091563A1 GPCR-Proteins and Nucleic Acids Encoding Sa
FILE REFERENCE: 15966-767
CURRENT APPLICATION NUMBER: US/09/832,522
PRIOR FILING DATE: 2001-04-11
PRIOR APPLICATION NUMBER: 60/195,994
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: 60/196,538
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: 60/220,644
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/264,851
PRIOR FILING DATE: 2001-01-29
PRIOR APPLICATION NUMBER: 60/199,964
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/268,567
PRIOR FILING DATE: 2001-02-14
PRIOR APPLICATION NUMBER: 60/199,955
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/259,641
PRIOR FILING DATE: 2001-01-04
PRIOR APPLICATION NUMBER: 60/200,176
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/199,948
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/199,956
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/218,995
PRIOR FILING DATE: 2000-07-17
NUMBER OF SEQ ID NOS: 112
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 74
LENGTH: 299
TYPE: PRT
ORGANISM: Homo sapiens
US-09-832-522-74

Query Match 24.1%; Score 49; DB 11; Length 299;
Best Local Similarity 33.3%; Pred. No. 63;
Matches 11; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Cy 2 NLGSRVFSKCKRLVFSWALVSATMAVTTV 34
Db 81 SLMGNGYISLVQICILQVFFIALASSEVALITV 113

RESULT 15
US-09-832-522-60
Sequence 60, Application US/09832522
Publication No. US20030091563A1
GENERAL INFORMATION:
```

APPLICANT: Shenoy, Suresh G  
APPLICANT: Gangoli, Esha A  
APPLICANT: Rastelli, Luca  
APPLICANT: Smithson, Glenda  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Vernet, Corine  
APPLICANT: Molenc, Adam R  
APPLICANT: Casman, Stacie J  
APPLICANT: Tchernev, Velizar T  
APPLICANT: Szekeres, Edward S  
APPLICANT: Grosse, William  
APPLICANT: Alsobrook, John P  
APPLICANT: Burgess, Catherine E  
TITLE OF INVENTION: No. US20030091563A1el GPCR-Proteins and Nucleic Acids Encoding Se  
FILE REFERENCE: 15966-767  
CURRENT FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/195,994  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: 60/196,538  
PRIOR FILING DATE: 2000-04-11  
PRIOR APPLICATION NUMBER: 60/220,644  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: 60/264,851  
PRIOR FILING DATE: 2001-01-29  
PRIOR APPLICATION NUMBER: 60/199,964  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/268,567  
PRIOR FILING DATE: 2001-02-14  
PRIOR APPLICATION NUMBER: 60/199,955  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/259,641  
PRIOR FILING DATE: 2001-01-04  
PRIOR APPLICATION NUMBER: 60/200,176  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/199,948  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/199,956  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/218,995  
PRIOR FILING DATE: 2000-07-17  
NUMBER OF SEQ ID NOS: 112  
SOFTWARE: PatentIn Ver. 2.1  
SEQ. ID NO 60  
LENGTH: 303  
TYPE: PRT  
ORGANISM: Pan troglodytes  
US-09-832-522-60

Query Match 24.1%; Score 49; DB 11; Length 303;  
Best Local Similarity 33.3%; Pred. No. 64;  
Matches 11; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

OY 2 NLGSRVFSKCRVLVFSMVALVSATMAVTV 34  
DB 83 SLMGNGYISLVQCILQVFFIALASSEVALTV 115

Search completed: January 5, 2004, 19:13:43  
Job time : 81.6648 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: January 5, 2004, 17:31:14 ; Search time 4.9895 Seconds  
(without alignments)  
828.882 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203

Sequence: 1 MNLGSRRVFSKRCRLVFS.....LVSAVTMTLENTALAR 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203	100.0	271	2	S36907
2	59	29.1	458	2	G72220
3	55.5	27.3	397	2	C53240
4	55.5	27.3	397	2	C39099
5	55	27.1	341	1	D69972
6	55	27.1	366	2	B69375
7	55	27.1	576	2	F75383
8	52.5	25.9	1019	2	T29623
9	52	25.6	215	2	SS7474
10	52	25.6	414	2	AG2996
11	52	25.6	467	2	B98287
12	51.5	25.4	206	2	S76279
13	51.5	25.4	221	2	S76798
14	51.5	25.4	728	2	B86093
15	51.5	25.4	728	2	G91245
16	51	25.1	658	2	B84664
17	51	25.1	838	2	D71492
18	51	25.1	1042	2	A97209
19	50.5	24.9	269	2	C91030
20	50.5	24.9	269	2	D85874
21	50.5	24.9	269	2	S08346
22	50.5	24.9	269	2	AD0804
23	50.5	24.9	598	2	T29878
24	50	24.6	243	2	A38901
25	50	24.6	271	2	J00980
26	50	24.6	888	2	C90595
27	49.5	24.4	130	2	G87422
28	49.5	24.4	375	2	AF1720
29	49.5	24.4	396	2	A39099

30	49.5	24.4	571	2	T36317	two-component sens
31	49.5	24.4	728	2	B65208	hypothetical 82.6
32	49	24.1	262	2	B81865	probable fimbrial
33	49	24.1	262	2	B81077	fimbrial assembly
34	49	24.1	476	1	CE4127	ADP-heptose syntha
35	49	24.1	1094	2	C59434	KIAL168 protein l
36	48.5	23.9	928	1	B38001	probable serine/ch
37	48	23.6	75	2	B36119	preoptic area regu
38	48	23.6	243	2	S07997	lipoprotein trar p
39	48	23.6	243	2	D84474	hypothetical prote
40	48	23.6	244	2	S01757	lipoprotein trar p
41	48	23.6	287	2	E75202	dipeptide abc tran
42	48	23.6	450	2	T37628	glycerol dehydroge
43	48	23.6	1204	2	C75015	probable pyrolysin
44	48	23.6	1759	2	T18868	myoblast cty prot
45	47.5	23.4	120	2	T18158	hypothetical prote

#### ALIGNMENTS

RESULT 1  
S36907  
mitogenic factor, 25K, precursor - Streptococcus pyogenes  
C:Species: Streptococcus pyogenes  
C:Date: 10-Dec-1993 #sequence revision 23-Feb-1996 #text\_change 15-Oct-1999  
C:Accession: S36907; S36908; S29188  
R:Yutendo, T.; Igarashi, H.; Hinuma, Y.; Yutendo, T.  
FEB8 Lett. 331, 187-192, 1993  
A:Title: Cloning, characterization and overexpression of a Streptococcus pyogenes gene e  
A:Reference number: S36907; MUID:94009636; PMID:8405402  
A:Accession: S36907  
A:Molecule type: DNA  
A:Residues: 1-271 <IM>  
A:Cross-references: EMBL:D13428; NID:9432369; PIDD:BA02693.1; PID:dl003198; PID:9432370  
A:Accession: S36908  
A:Molecule type: protein  
A:Residues: 44-64 <IM>  
A:Yutendo, T.; Murai, H.; Gonzalez, J.; Takao, T.; Shimonishi, Y.; Takeda, Y.; Igarashi,  
FEB8 Lett. 308, 30-34, 1992  
A:Title: A new type of mitogenic factor produced by Streptococcus pyogenes.  
A:Reference number: S29188; MUID:92354761; PMID:1644200  
A:Accession: S29188  
A:Molecule type: protein  
A:Residues: 44-64 <YUT>  
F:1-43/Domain: signal sequence #status predicted <SIG>  
F:44-271/Product: mitogenic factor, 25K #status experimental <MAT>  
Query Match 100.0%; Score 203; DB 2; Length 271;  
Best local similarity 100.0%; Pred. No. 7e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MNLGSRRVFSKRCRLVFSNALVSATMAVTMTLENTALAR 43  
DB 1 MNLGSRRVFSKRCRLVFSNALVSATMAVTMTLENTALAR 43

RESULT 2  
G72220  
conserved hypothetical protein - Thermotoga maritima (strain MS88)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72220  
R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316; PMID:10360571  
A:Accession: G72220  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-458 <ARN>

A:Cross-references: GB:AE001810; GB:AE000512; NID:g4982271; PIDN:AA036768.1; PID:g4982271  
 A:Experimental source: strain M8B8  
 C:Gene: TM1701  
 C:Superfamily: Mechanococcus jannaschii conserved hypothetical protein MJ0709

Query Match 29.1%; Score 59; DB 2; Length 458;  
 Best Local Similarity 44.4%; Pred. No. 2.1;  
 Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 9 VPSKKCRVKSMAVLSATMAVTTVT 35  
 DB 276 VFTSAMRVNFGTVPLIGMAVAVTSVT 302

RESULT 3  
 CS3240  
 allergen Amb a 1.3 precursor - common ragweed  
 C:Species: Ambrosia artemisiifolia (common ragweed)  
 C>Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 18-Sep-1998  
 C:Accession: CS3240  
 R:Griffith, I.J.; Pollack, J.; Klapper, D.G.; Rogers, B.L.; Nault, A.K.  
 Int. Arch. Allergy Appl. Immunol. 96, 296-304, 1991  
 A:Title: Sequence polymorphism of Amb a 1 and Amb a II, the major allergens in Ambrosia  
 A:Reference number: A53240; MUID:92234570; PMID:1809687  
 A:Accession: CS3240  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-397 <GR1>  
 A:Cross-references: GB:M80560  
 C:Superfamily: pectate lyase LATS9  
 C:Keywords: pollen

Query Match 27.3%; Score 55.5; DB 2; Length 397;  
 Best Local Similarity 36.8%; Pred. No. 5.6;  
 Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNLLGSRVPSKKCRVKSMAVLSATMAVTTVTLEN 38  
 DB 198 INVAGSSQIWDHCSLSK-SFDGLVDVTLGSTRHTTISN 234

RESULT 4  
 CS3099  
 allergen Amb a 1.3 - common ragweed  
 C:Species: Ambrosia artemisiifolia (common ragweed)  
 C>Date: 27-Nov-1991 #sequence\_revision 03-Apr-1992 #text\_change 29-Sep-1999  
 C:Accession: CS3099  
 R:Rathar, T.; Griffith, I.J.; Kuo, M.; Bond, J.F.; Rogers, B.L.; Klapper, D.G.  
 J. Biol. Chem. 266, 1229-1236, 1991  
 A:Title: Cloning of Amb a 1 (antigen E), the major allergen family of short ragweed poll  
 A:Reference number: A39099; MUID:91093235; PMID:1702434  
 A:Accession: CS3099  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-397 <RAF>  
 A:Cross-references: GB:M2961; NID:g166440; PIDN:AA32668.1; PID:g166441  
 C:Superfamily: pectate lyase LATS9  
 C:Keywords: pollen

Query Match 27.3%; Score 55.5; DB 2; Length 397;  
 Best Local Similarity 36.8%; Pred. No. 5.6;  
 Matches 14; Conservative 7; Mismatches 16; Indels 1; Gaps 1;

QY 1 MNLLGSRVPSKKCRVKSMAVLSATMAVTTVTLEN 38  
 DB 198 INVAGSSQIWDHCSLSK-SFDGLVDVTLGSTRHTTISN 234

RESULT 5  
 D69972  
 opine catabolism homolog yrbB - Bacillus subtilis  
 C:Species: Bacillus subtilis

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: D69972

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berti  
 C: Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
 A: Ehrlich, S.D.; Emmerich, P.T.; Entian, K.D.; Ertling, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gall  
 lech, J.; Harwood, C.R.; Hensut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F  
 Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel  
 Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadala, Y.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror  
 T.; Winters, P.; Wipac, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumelein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; MUID:98044033; PMID:9384377  
 A:Accession: D69972  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-341 <KUN>  
 A:Cross-references: GB:Z99118; GB:AL009126; NID:g2635200; PIDN:CB14737.1; PID:g2635242  
 A:Experimental source: strain 168  
 C:Gene: yrbB  
 C:Superfamily: Streptomyces griseus strI protein

Query Match 27.1%; Score 55; DB 1; Length 341;  
 Best Local Similarity 38.1%; Pred. No. 5.7;  
 Matches 16; Conservative 6; Mismatches 18; Indels 2; Gaps 1;

QY 3 LLGSR--RVPSKKCRVKSMAVLSATMAVTTVTLENTALA 42  
 DB 186 IMGSEVTEVYAKGALVNPSPAEIGDIDTAVITLTFENGAMA 227

RESULT 6  
 B69375  
 ABC transporter, ATP-binding protein homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: B69375  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Uettermann, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: B69375  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-366 <KLE>  
 A:Cross-references: GB:AE001034; GB:AE000782; NID:g2689357; PIDN:AA090233.1; PID:g264958

Query Match 27.1%; Score 55; DB 2; Length 366;  
 Best Local Similarity 29.7%; Pred. No. 6.1;  
 Matches 11; Conservative 10; Mismatches 16; Indels 0; Gaps 0;

QY 7 RRVPSKKCRVKSMAVLSATMAVTTVTLENTALAR 43  
 DB 155 RFVFTTKSRLLDFVAPALIGVTVQMLAILTSSSIAR 191

RESULT 7  
 F75383  
 ribS protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: F75383  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.U.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Frazer, C.M.  
 Science 286, 1571-1577, 1999  
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A|Reference number: A75250; MUID:20036896; PMID:10567266  
 A|Accession: F75383  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-576 <WHI>  
 A|Cross-references: GB:AE001998; GB:AE000513; NID:g6459302; PIDN:AAF1110.1; PID:g645930  
 A|Experimental source: strain R1  
 C|Genetics:  
 A|Gene: DR1550  
 A|Map position: 1

Query Match 27.1%; Score 55; DB 2; Length 576;  
 Best Local Similarity 35.7%; Pred. No. 9.5; 19; Indels 0; Gaps 0;  
 Matches 15; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MNLGSRVPSKRCRLVKSVALVSATMAVTTVLTAL 42  
 Db 485 MRLGATRSFTRMPLRLGLVGVAVLATVATAMGVQVLA 526

RESULT 8  
 T29623  
 hypochlorite protein T05E8.3 - Caenorhabditis elegans  
 C|Species: Caenorhabditis elegans  
 C|Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000  
 C|Accession: T29623  
 R|Wu, X.; Le, T.T.  
 A|Submitted to the EMBL Data Library, April 1997  
 A|Description: The sequence of C. elegans cosmid T05E8.  
 A|Reference number: Z20654  
 A|Accession: T29623  
 A|Status: preliminary; translated from GB/EMBL/DBJ  
 A|Molecule type: DNA  
 A|Residues: 1-1019 <WUX>  
 A|Cross-references: EMBL:U97014; PIDN:AA852427.1; GSPDB:GN00019; CESP:T05E8.3  
 A|Experimental source: strain Bristol N2; clone T05E8  
 C|Genetics:  
 A|Gene: CESP:T05E8.3  
 A|Map position: 1  
 A|Intons: 26/1; 110/3; 196/3; 264/2; 339/1; 491/2; 556/3; 600/3; 633/2; 671/2; 731/3; 7

Query Match 25.9%; Score 52.5; DB 2; Length 1019;  
 Best Local Similarity 29.5%; Pred. No. 38;  
 Matches 13; Conservative 11; Mismatches 5; Indels 15; Gaps 2;

Qy 8 RVFSKRCR-----LVKFSMA-----LVSATMAVTTVLT 36  
 Db 575 RFRSRRCHPSCGSLKILKFTPLSGARKVITINIMETSVTI 618

RESULT 9  
 S57474  
 GTP-binding protein - garden pea  
 N|Alternate names: small GTP-binding protein  
 C|Species: Pisum sativum (garden pea)  
 C|Date: 10-Oct-1995 #sequence\_revision 17-Nov-1995 #text\_change 24-May-2001  
 C|Accession: S57474  
 R|Rohli, T.; Caliebe, A.; Seedorf, M.; Soll, J.  
 A|Submitted to the EMBL Data Library, June 1995  
 A|Description: Pea mRNA for GTP-binding protein related to the Ypt/Rab-subfamily.  
 A|Reference number: S57462  
 A|Accession: S57474  
 A|Molecule type: mRNA  
 A|Residues: 1-215 <ROE>  
 A|Cross-references: EMBL:Z49899; NID:8871513; PIDN:CAA0079.1; PID:8871514  
 C|Superfamily: ras transforming protein; translation elongation factor Tu homology  
 C|Keywords: GTP binding; nucleotide binding; P-loop  
 F|16-131/Domain: translation elongation factor Tu homology <ETU>  
 F|22-29/Region: nucleotide-binding motif A (P-loop)

F|128-131/Region: GTP-binding NKXD motif  
 F|159-161/Region: GTP-binding SAK/L motif

Query Match 25.6%; Score 52; DB 2; Length 215;  
 Best Local Similarity 31.0%; Pred. No. 9.6; 13; Indels 8; Gaps 2;  
 Matches 13; Conservative 8; Mismatches 13; Indels 8; Gaps 2;

Qy 3 LLGSRVPSKRCRLVKS-----SMVALVSATMAVTTVLT 37  
 Db 20 LIGRRV-GRGCLLRFDGSGFTSTFATIGIDFIRITIED 60

RESULT 10  
 AG2996  
 transcription regulator Atu3577 [imported] - Agrobacterium tumefaciens (strain C58, Dupo.  
 C|Species: Agrobacterium tumefaciens  
 C|Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C|Accession: AG2996  
 R|Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCall.  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A|Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Kreppan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A|Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A|Reference number: AB2577; MUID:21608550; PMID:11743193  
 A|Accession: AG2996  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-414 <KUR>  
 A|Cross-references: GB:AE006889; PIDN:AAU44389.1; PID:g17741986; GSPDB:GN00187  
 A|Experimental source: strain C58 (Dupont)  
 C|Genetics:  
 A|Gene: Atu3577  
 A|Map position: linear chromosome

Query Match 25.6%; Score 52; DB 2; Length 414;  
 Best Local Similarity 29.3%; Pred. No. 18;  
 Matches 12; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MNLGSRVPSKRCRLVKSVALVSATMAVTTVLTAL 41  
 Db 20 LSVMAALRLYGLSLHTDMSFTGLASATVSVITDLERTGL 60

RESULT 11  
 B98287  
 probable transcription regulator [imported] - Agrobacterium tumefaciens (strain C58, Cer.  
 C|Species: Agrobacterium tumefaciens  
 C|Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C|Accession: B98287  
 R|Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A|Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A|Reference number: A97359; MUID:21608551; PMID:11743194  
 A|Accession: B98287  
 A|Status: preliminary  
 A|Molecule type: DNA  
 A|Residues: 1-467 <KUR>  
 A|Cross-references: GB:AE007870; PIDN:AAK89820.1; PID:g15159752; GSPDB:GN00170  
 C|Genetics:  
 A|Gene: AGR\_L\_2504  
 A|Map position: linear chromosome

Query Match 25.6%; Score 52; DB 2; Length 467;  
 Best Local Similarity 29.3%; Pred. No. 21;  
 Matches 12; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MNLGSRVPSKRCRLVKSVALVSATMAVTTVLTAL 41  
 Db 73 LSVMAALRLYGLSLHTDMSFTGLASATVSVITDLERTGL 113

## RESULT 12

S76279

A:Species: Synechocystis sp. (strain PCC 6803)

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S76279

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76279

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-206 &lt;KAN&gt;

A:Cross-references: EMBL:D64000; GB:AB001339; NID:g1001484; PIDN:BAI0131.1; PID:d101078

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match

Best Local Similarity 25.4%; Score 51.5; DB 2; Length 206;

Matches 13; Conservative 10; Mismatches 17; Indels 1; Gaps 1;

QY 3 ILGSRVRSKR-CRLVKSMAVLSATMAVTTVLTENTALA 42

Db 62 LFGSKKLGGQVLTQVQLFTPLVGAMMAIVARLLTSELA 102

## RESULT 13

S76798

A:Species: Synechocystis sp. (strain PCC 6803)

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S76798

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-116, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S76798

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-221 &lt;KAN&gt;

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAI18710.1; PID:d101944

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 25.4%; Score 51.5; DB 2; Length 221;

Best Local Similarity 41.9%; Pred. No. 12;

Matches 13; Conservative 10; Mismatches 3; Indels 5; Gaps 2;

QY 12 KCCRLVKSMAVLSATMAVTTVLTENTALA 42

Db 24 KRCRL--WLSLTLALVMASTLT--ALA 49

## RESULT 14

E86093

A:Species: Escherichia coli (strain O157:H7, substr

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: E86093

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Miller, L.; Grobbeck, E.J.; Davis, N.W.; Linn, A.; Dimmlanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E86093

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-728 &lt;STO&gt;

A:Cross-references: GB:AE005174; NID:g12518947; PIDN:AAG59209.1; GSPDB:GN00145; UMGPR:256

A:Experimental source: strain O157:H7, substrain EDL933

A:Genetics:

A:Gene: atp

Query Match 25.4%; Score 51.5; DB 2; Length 728;

Best Local Similarity 34.2%; Pred. No. 37;

Matches 13; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 7 RRVSKRCRLVKSMAVLSATMAVTTVLT-ENTALAR 43

Db 60 RHIFSNKDFVIFKFSIVLQADKKEITLNNKENTTLTQ 97

## RESULT 15

G91245

A:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C:Accession: G91245

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Gasawara, N.; Yasunaga, T.; Kunihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: G91245

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-728 &lt;HAY&gt;

A:Cross-references: GB:BA000007; PIDN:BAH38358.1; PID:g13364411; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

A:Genetics:

A:Gene: EC84935

Query Match 25.4%; Score 51.5; DB 2; Length 728;

Best Local Similarity 34.2%; Pred. No. 37;

Matches 13; Conservative 8; Mismatches 16; Indels 1; Gaps 1;

QY 7 RRVSKRCRLVKSMAVLSATMAVTTVLT-ENTALAR 43

Db 60 RHIFSNKDFVIFKFSIVLQADKKEITLNNKENTTLTQ 97

Search completed: January 5, 2004, 18:42:46  
Job time : 10.989 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:46:12 ; Search time 3.64273 Seconds  
(without alignments)  
555.120 Million cell updates/sec

Title: US-08-482-785-1  
Perfect score: 203  
Sequence: 1 MNLGSRVRSKCKLVKFS.....LVSATMAVTTLTALALR 43

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues  
Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55.5	27.3	397	1 MP13_AMBAR	P27761 ambrosia ar
2	55	27.1	341	1 YRBE_BACCU	O05389 escherichia su
3	50.5	24.9	269	1 YFCA_ECOLI	P14008 escherichia
4	50.5	24.9	458	1 ARLY_HELMO	O08dus heliobacilli
5	50	24.6	243	1 TRT2_ECOLI	P32885 escherichia
6	49.5	24.4	396	1 MP11_AMBAR	P27759 ambrosia ar
7	49.5	24.4	728	1 ARPA_ECOLI	P23325 escherichia
8	49	24.1	321	1 OSU1_HUMAN	O949f5 homo sapien
9	49	24.1	430	1 MURA_RHIME	O92827 rhizobium m
10	49	24.1	476	1 RFAE_HAEIN	O05074 haemophilus
11	49	24.1	600	1 BIR3_MOUSE	O08863 mus musculu
12	49	24.1	1083	1 KG88_HUMAN	O9c0h5 homo sapien
13	48.5	23.9	928	1 KKRI_YEAST	P36003 saccharomyc
14	48	23.6	101	1 KG88_RAT	P18890 rattus norv
15	48	23.6	243	1 TRT4_ECOLI	P15177 escherichia
16	48	23.6	244	1 TRT1_ECOLI	P13979 escherichia
17	48	23.6	661	1 NKX2_HUMAN	O9u140 homo sapien
18	48	23.6	1097	1 KG88_MOUSE	P53281 mus musculu
19	47.5	23.4	332	1 SRG3_CAEBL	P45572 caenorhabdi
20	47	23.2	457	1 YEBQ_ECOLI	P76269 escherichia
21	46.5	22.9	659	1 CYOB_BUCAP	O84994 buchnera ap
22	46	22.7	243	1 TRT1_SALTY	P22107 salmonella
23	46	22.7	302	1 Y016_RICPR	O92ec7 rickettsia
24	46	22.7	424	1 MURA_HAEIN	P45025 haemophilus
25	46	22.7	430	1 MUAI_RHIO	O98965 rhizobium 1
26	46	22.7	469	1 YLJ1_CAEBL	P34366 caenorhabdi
27	46	22.7	1391	1 LYS2_CANAL	O12572 candida alb
28	45.5	22.4	230	1 TOLQ_ECOLI	P05828 escherichia
29	45.5	22.4	294	1 MENA_MYCLE	O07134 mycobacteri
30	45.5	22.4	635	1 VF12_VACCC	P21053 vaccinia vi
31	45	22.2	374	1 PACF_HUMAN	O9np18 homo sapien
32	45	22.2	513	1 LEU1_LACIA	O02141 lactococcus
33	45	22.2	552	1 GPW1_LEPIN	P59173 leptospira

## ALIGNMENTS

RESULT 1	ID	MP13_AMBAR	STANDARD	PRT	397 AA.
AC	P27761				
DT	01-AUG-1992	(Rel. 23, Created)			
DT	01-AUG-1992	(Rel. 23, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Pollen allergen Amb a 1.3 precursor (Antigen E) (Antigen Amb a I).				
OS	Ambrosia artemisiifolia (short ragweed).				
OC	Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;				
OC	Spermatophyta: Magnoliophyta; eudicotyledons; core eudicots;				
OC	Asteridae; campanulids; Asterales; Asteraceae; Asteroideae;				
OC	Heliantheae; Ambrosia.				
OX	NCBI_TaxID=4212;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISUE=Pollen;				
RX	MEDLINE=91093235; PubMed=1702434;				
RA	Ratner T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,				
RA	Klapper D.G.;				
RT	"Cloning of Amb a I (antigen E), the major allergen family of short				
RT	ragweed pollen.";				
RL	J. Biol. Chem. 266:1229-1236(1991).				
RN	[2]				
RP	SEQUENCE FROM N.A., AND VARIANTS.				
RC	TISUE=Pollen;				
RX	MEDLINE=92234570; PubMed=1809687;				
RA	Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.;				
RT	"Sequence polymorphism of Amb a I and Amb a II, the major allergens				
RT	in Ambrosia artemisiifolia (short ragweed).";				
RL	Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).				
CC	-1- SUBUNIT: Monomer.				
CC	-1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.				
CC	-1- PTM: The N-terminus is blocked.				
CC	-1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.				
CC	-1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.				
CC	AMB A I/AMB A II/CRY J I SUBFAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.1ab-stb.ch/announce/">http://www.1ab-stb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@1ab-stb.ch">license@1ab-stb.ch</a> ).				
CC	-----				
DR	EMBL; M62961; AAA32668.1; -				
DR	EMBL; M80560; AAA32669.1; ALT_SEQ.				
DR	PIR; C39099; C39099.				
DR	InterPro; IPR002022; Amb allergen.				
DR	Pfam; PF00544; pec_lyase_1.				
DR	PRINTS; PR00807; AMBALDERGEN.				
DR	SMART; SM00656; Amb_a1; 1.				
KW	Antigen; Allergen; Signal; Multigene family; Polymorphism.				
FT	SIGNAL	1	25	POTENTIAL.	
FT	CHAIN	26	397	POLLEN ALLERGEN AMB A 1.3.	

ST	VARIANT#	48	48	L > Y
QC	SEQUENCE	397 AA;	42928 MM;	C8D8A1257590DD0A CRC64;
	Query Match		27.3%;	Score 55.5; DB 1; Length 397;
	Best Local Similarity	36.8%;	Pred. No. 2;	
	Matches 14;	Conservative	7;	Mismatches 16; Indels 1; Gaps
Qy	1	MLNLGSRVRSFKKCRNLVKEFMVALVYSATMAVTVTLEN	38	
Db	198	INWAGSSQIWIIDHCISLKR-SFDGLVDTLIGSTHTVTLSN	234	
RESULT 2				
YRBE_BACSU				
ID	YRBE_BACSU	STANDARD;	PRT;	341 AA.
AC	005389;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, last sequence update)			
DT	28-FEB-2003 (Rel. 41, last annotation update)			
DE	Hypothetical oxidoreductase yrbE (EC 1.-.-.-).			
GN	YRBE.			
OS	Bacillus subtilis.			
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.			
NCBI	TaxID=1423;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=168;			
RC	SEQUENCE FROM N.A.			
RX	MEDLINE=98044033; PubMed=9384377;			
RA	Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,			
RA	Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,			
RA	Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,			
RA	Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,			
RA	Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,			
RA	Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,			
RA	Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,			
RA	Fritz C., Fujita M., Fujita Y., Funas S., Galizzi A., Galleron N.,			
RA	Ghm S.Y., Glaeser P., Goffeau A., Golligly E.J., Grandi G.,			
RA	Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,			
RA	Hilbert H., Holappell S., Hosono S., Hullo M.F., Itaya M., Jones L.,			
RA	Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,			
RA	Kobayashi Y., Koelter P., Konigstein G., Krogh S., Kumano M.,			
RA	Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,			
RA	Lee S.W., Levine A., Liu H., Maeda S., Meuel C., Medigue C.,			
RA	Medina N., Meliando R.P., Mizuno M., Moesti D., Nakai S., Noback M.,			
RA	Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,			
RA	Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,			
RA	Prescan E., Pujic P., Punnelle B., Rapoport G., Rey M., Reynolds S.,			
RA	Rieger M., Rivolta C., Rooha B., Roche B., Rose M., Sadate Y.,			
RA	Sato T., Scallan E., Schlich S., Schreiber R., Scoffone F.,			
RA	Sekiguchi J., Sekowska A.A., Serrif S.J., Serrif P., Shin B.S., Soldo B.,			
RA	Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,			
RA	Takenuchi M., Tanakoshi A., Tanaka T., Terpsita P., Tognoni A.,			
RA	Toato V., Uchiyama S., Vandenbol M., Vamier F., Vassarotti A.,			
RA	Vitari A., Wandut R., Wedler B., Wedler H., Weltensager T.,			
RA	Winters P., Wipit A., Yamamoto H., Yamane K., Yamamoto K., Yata K.,			
RA	Yoshida K., Yoshikawa H.F., Zumestein E., Yoshikawa H., Danchin A.,			
RT	"The complete genome sequence of the Gram-positive bacterium Bacillus			
RT	subtilis.";			
RL	Nature 390:249-256(1997).			
RN	[3]			
RN	SEQUENCE OF 252-341 FROM N.A.			
RC	STRAIN=168 / PY79;			
RX	MEDLINE=97177783; PubMed=9025289;			
RA	Gomez M., Cutting S.M.;			
RT	"BofC encodes a putative forespore regulator of the Bacillus subtilis			
RT	sigma K checkpoint.";			

```

RL Microbiology 143:157-170(1997).
-i SIMILARITY: BELONGS TO THE GEO_IDH/MOCA FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by end-for-commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
or send an email to license@isb-sib.ch.)
-----
DR EMBL; Y15896; CAB75327.1; -.
DR EMBL; Z99118; CAB14737.1; -.
DR EMBL; X93081; CA63619.1; -.
DR PIR; D69972; D69972.
DR Subtilisin; BGI2280; yrbE.
DR InterPro; IPR000683; GEO_IDH_MoCA_.
DR InterPro; IPR004104; GEO_IDH_MoCA_C_.
DR Pfam; PF01408; GEO_IDH_MoCA_1.
DR Pfam; PF02894; GEO_IDH_MoCA_C_1.
DM Hypothetical protein; Oxidoreductase; Complete proteome.
SQ SEQUENCE 341 AA; 3778 MW; E3023CD874F90C5F CRC64;
Query Match      27.1%; Score 55; DB 1; Length 341;
Best Local Similarity   38.1%; Pred. No. 2;
Matches    16; Conservative     6; Mismatches    18; Indels       2; Gaps        1;
OY          3 LIGSR--RVPSSKCRLYKFNSVALVSATMAVTVLTLENTALA 42
Db           186 IMGSEVFTEVKAGALVNPSFAELGIDDTAVITLFENGAWA 227
::||| ::||| ||| | | | | | :|
RESULT 3
YCFC ECOLI STANDARD; PROT; 269 AA.
AC PI4008;
AD 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ycfcA.
GN YCFC OR B2327 OR C2873 OR Z3590 OR ECS3211.
OS Escherichia coli.
OS Escherichia coli O6, and
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=X12;
RX MEDLINE=90251161; PubMed=2187143;
RA Keck W., van Leeuwen A.M., Huber M., Goodell E.W.;
RT "Cloning and characterization of mepA, the structural gene of the
penicillin-insensitive murcin endopeptidase from Escherichia coli.";
RL Mol. Microbiol. 4:209-219(1990).
[2]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=X12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirtpatrick H.A., Goeden M.A., Rose D.J.,
Maun B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=X12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Aliba H., Baba T., Hayashi K., Inada T., Isoe K.,
RA Ichio T., Kimura S., Kitagawa M., Makino K., Miki T., Minohashi N.,
RA Mizubuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Ohshima T., Oyama S., Saito N., Sampei G., Saton Y., Sivaenudaram S.,
```

RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horiiuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*  
 RT - K12 genome corresponding to 50,0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O6:H1 / CFT073 / ATCC 700928;  
 RC MEDLINE=2238234; PubMed=12471157;  
 RA Welch R.A., Burland V., Plunkett G., III, Redford P., Roesch P.,  
 RA Raebro D., Buckles E.L., Lion S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 RT of uropathogenic *Escherichia coli*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RC MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Posfai G., Hackett J., Link S., Boutin A., Shao Y., Miller L.,  
 RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 [6]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=O157:H7 / RIMD 0509952;  
 RC MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohnishi E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic *Escherichia coli*  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
 CC -1- SIMILARITY: STRONG, TO H.INFLUENZA H10198.  
 CC -1- SIMILARITY: TO M.LEPRAE U1177B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL; X16909; CAAC7483.1; -;  
 DR EMBL; AE000321; AAC75387.1; -;  
 DR EMBL; D90863; BAA16183.1; -;  
 DR EMBL; AE016764; AAN61323.1; -;  
 DR EMBL; AE005464; AAG57455.1; -;  
 DR EMBL; AF002561; BAB36634.1; -;  
 DR PIR; C91030; C91030.  
 DR PIR; D85874; D85874.  
 DR PIR; S08346; S08346.  
 DR EcoGene; EG11444; yfca.  
 DR InterPro; IPR002781; DUF81.  
 DR Pfam; PF01925; DUF81; 1.  
 DR Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 8 28 POTENTIAL.  
 FT TRANSMEM 31 51 POTENTIAL.  
 FT TRANSMEM 85 105 POTENTIAL.  
 FT TRANSMEM 112 132 POTENTIAL.  
 FT TRANSMEM 157 177 POTENTIAL.  
 FT TRANSMEM 198 218 POTENTIAL.  
 SQ SEQUENCE 269 AA; 28643 MW; B640B0BFF1BF7ED CRC64;

Query Match 24.9%; Score 50.5; DB 1; Length 269;  
 Best Local Similarity 44.4%; Pred. No. 7.1;  
 Matches 12; Conservative 7; Mismatches 7; Indels 1; Gaps 1;  
 QY 4 LGSRVFSKCRIVKFSMVALVSATMA 30  
 DB 225 MGSRLVLSKQKLR-PMIVISAVMS 250  
 RESULT 4  
 ID ARLY\_HELMO STANDARD; PRT; 458 AA.  
 AC Q8GD05;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Argininosuccinate lyase (EC 4.3.2.1) (Argininosuccinase) (ASAL)  
 DE (Fragment).  
 OS Hellobacillus mobilis.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Hellobacteriaceae;  
 OC Hellobacillus.  
 OX NCBI\_TaxID=28064;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RH0002707;  
 RC MEDLINE=22337798; PubMed=12446909;  
 RA Raymond J., Zhaxybayeva O., Gogarten J.P., Gerdes S.Y.,  
 RA Blankenship R.E.;  
 RT "Whole-genome analysis of photosynthetic prokaryotes.";  
 RL Science 298:1616-1620(2002).  
 CC -1- CATALYTIC ACTIVITY: N-(L-arginino)succinate = fumarate + L-  
 CC arginine.  
 CC -1- PATHWAY: Arginine biosynthesis; eighth (last) step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (probable).  
 CC -1- SIMILARITY: Belongs to the lyase I family. Argininosuccinate lyase  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC  
 DR EMBL; AY142879; AAN67483.1; -;  
 DR HAMAP; MF\_00006; -; 1.  
 DR InterPro; IPR003031; D\_cryta11n.  
 DR InterPro; IPR000362; Fumarate\_lyase.  
 DR Pfam; PF00206; Lyase\_1; 1.  
 DR PRINTS; PR00149; FUMKATELYASE.  
 DR TIGRFAMs; TIGR00838; argH; 1.  
 DR PROSITE; PS00163; FUMARATE\_LYASES; 1.  
 DR Arginine biosynthesis; Lyase.  
 FT NON TER 458 458  
 SQ SEQUENCE 458 AA; 51030 MW; 8A7EC9F7E41C8AC CRC64;  
 Query Match 24.9%; Score 50.5; DB 1; Length 458;  
 Best Local Similarity 34.0%; Pred. No. 12;  
 Matches 16; Conservative 11; Mismatches 15; Indels 5; Gaps 2;  
 QY 2 NLLGSRVFSKRC-RIV-----KFSMVALVSATMAVTTLTENTLAR 43  
 DB 169 HLAVVQMFGRDCSRIVDCARFNIPLSGSALGATTFPDRMAVQ 215  
 RESULT 5  
 ID TRT2\_ECOLI STANDARD; PRT; 243 AA.  
 AC P32885; P07177;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1988 (Rel. 07, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

```

OX NCBI_TaxID=4212;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=PolLEN;
RX MEDLINE=91093235; PubMed=1702434;
RA Ratner T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
RN Klapper D.G.;
RT "Cloning of Amb A I (antigen E), the major allergen family of short
RL ragweed pollen.",
RN J. Biol. Chem. 266:1229-1236(1991).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=PolLEN;
RX MEDLINE=92234570; PubMed=1809687;
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Mauls A.K.;
RT "Sequence polymorphism of Amb A I and Amb A II, the major allergens
RN in Ambrosia artemisiifolia (short ragweed).";
RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC -1 SUBUNIT: Monomer.
CC -1 TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1 PM: The N-terminus is blocked.
CC -1 DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1 SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A I/AMB A II/CRY J I SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M6316; -; NOT ANNOTATED_CDS.
DR EMBL; M60558; AA33265.1; -.
DR PIR; A39099; A39099.
DR InterPro; IPR002022; Amb_allergen.
DR Pfam; PF00544; pec_lyase; 1.
DR PRINTS; PRO0807; AMBALLERGEN.
DR SMART; SM00656; Amb_11; 1.
KW Antigen; Allergen; Signal; Multigene family; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 396 POTENTIAL.
FT VARIANT 92 92 POLLEN ALLERGEN AMB A 1.1.
FT SEQUENCE 396 AA; 42709 MW; 0CE7DDECB2B8841D CRC64;
SQ
Query Match 24.4%; Score 49.5; DB 1; Length 396;
Best Local Similarity 30.8%; Pred. No. 15;
Matches 12; Conservative 11; Mismatches 15; Indels 1; Gaps 1;
QY 1 MNLIGSRVFSKCRVLVKSVALVSAVTAVTTLTNT 39
ID :: :: :: :: :: :: :: :: :: :: :: ::
ID ARPA_ECOLI STANDARD; PRT; 728 AA.
AC P23325; P76781;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DN Ankyrin-repeat protein A (Ankyrin-like regulatory protein).
DB ARPA OR ARP OR B4017.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,

```

```

RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RL region from 89.2 to 92.8 minutes.";
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [2]
RP SEQUENCE OF 108-728 FROM N.A.
RX MEDLINE=91138981; PubMed=1995429;
RA Galliner A., Bleicher F., Negre D., Perriere G., Duclos B.,
RT Cozzone A.J., Cortay J.-C.;
RT "Primary structure of the intergenic region between acek and iclr in
the Escherichia coli chromosome.";
RL Gene 97:149-150(1991).
RN [3]
RP IDENTIFICATION OF ANKYRIN-REPEATS.
RX MEDLINE=94285214; PubMed=8014990;
RA Newwald A.P., Green P.;
RT "Detecting patterns in protein sequences.";
RL J. Mol. Biol. 239:698-712(1994).
RN [4]
RP CHARACTERIZATION.
RA Kumari S., Beatty C., Wolfe A.J.;
RL unpublished observations (Aug-1996).
CC -1- FUNCTION: REGULATOR OF ACS SYNTHESIS.
CC -1- SIMILARITY: Contains 5 ANK repeats.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U00006; AAC3111.1; -
DR EMBL; AE000474; AAC76987.1; -
DR EMBL; M63497; AAA73004.1; -
DR PIR; H65208; H65208.
DR EcoGene; EG1208; arpa.
DR InterPro; IPR002110; ANK.
DR Pfam; PF00023; ank; 2.
DR SMART; SM00248; ANK; 1.
DR PROSITE; PS50088; ANK_REPEAT; FALSE NEG.
DR PROSITE; PS50297; ANK_REPEAT_REGION; FALSE NEG.
KW Repeat; ANK repeat; Complete proteome.
FT REPEAT 381 410 ANK 1.
FT REPEAT 429 458 ANK 2.
FT REPEAT 477 506 ANK 3.
FT REPEAT 525 554 ANK 4.
FT REPEAT 573 602 ANK 5.
FT CONFLICT 124 124 N -> D (IN REF. 2).
FT CONFLICT 282 282 T -> S (IN REF. 1; AAC76987).
FT CONFLICT 701 728 GFTDPRYIAEKNYEALUKASPTIVR -> TOKSISPYR
SQ SEQUENCE 728 AA; 82612 MW; 02577283E823928 CRC64;
Query Match 24.4%; Score 49.5; DB 1; Length 728;
Best Local Similarity 34.2%; Pred. No. 28;
Matches 13; Conservative 7; Mismatches 17; Indels 1; Gaps 1;

```

```

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: PUTATIVE ODORANT RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AL096770; CAB65799.1; -
DR Genem; HGNC:13971; OR5U1.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN_RECPT_F1_1; 1.
DR PROSITE; PS50262; G-PROTEIN_RECPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KM Multigene family; Olfaction.
FT DOMAIN 1 23 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 24 47 1 (POTENTIAL).
FT DOMAIN 48 55 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 56 77 2 (POTENTIAL).
FT DOMAIN 78 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 118 3 (POTENTIAL).
FT DOMAIN 119 137 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 138 156 4 (POTENTIAL).
FT DOMAIN 157 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 217 5 (POTENTIAL).
FT DOMAIN 218 234 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 235 257 6 (POTENTIAL).
FT DOMAIN 258 270 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 271 290 7 (POTENTIAL).
FT DOMAIN 291 321 CYTOPLASMIC (POTENTIAL).
FT DISULFD 95 187 BY SIMILARITY.
FT CARBOHYD 3 3 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 321 AA; 35891 MW; 69B33B0A1F49772 CRC64;
Query Match 24.1%; Score 49; DB 1; Length 321;
Best Local Similarity 33.3%; Pred. No. 14;
Matches 11; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

```

```

OY 2 NLLGSRVFSKCKLVKFSVALVSATMAVTV 34
DB 83 SLMGVYISLVQICILQVFFIALSSVALITV 115

```

```

RESULT 9
MURA_RHIME
ID MURA_RHIME STANDARD; PRT; 430 AA.
AC Q92827;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
DE (Enoylpyruvate transferase) (UDP-N-acetylglucosamine enolpyruvyl
transferase) (EPT).
GN MURA OR R00609 OR SMC02305.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxId=382;
RN [1]

```

SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RA MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Botte G., Ampe F., Batut J.,  
 RA Boistard P., Becker A., Boutry M., Cadieu B., Dreano S., Gloux S.,  
 RA Goerke T., Goffeau A., Kahn D., Kiss E., Lelaine V., Maury D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rampeger U.,  
 RA Renard C., Thebault P., Vandebol M., Weidner S., Gallibert F.,  
 RA "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021".  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: Cell wall formation. Adds enolpyruvyl to UDP-N-  
 CC acetylglucosamine (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + UDP-N-acetyl-D-  
 CC glucosamine = phosphate + UDP-N-acetyl-3-(1-carboxyvinyl)-D-  
 CC glucosamine.  
 CC -1- PATHWAY: Peptidoglycan biosynthesis; first step.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: Belongs to the BPS synthase family. MurA subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL591784; CAC45181.1; -.  
 DR HAMAP: MF\_00111; -. 1  
 DR InterPro: IPR005750; AcgH\_Tran\_MurA.  
 DR InterPro: IPR001986; BPS\_synthase.  
 DR Pfam: PF00275; BPS\_synthase; 1.  
 DR Prodom: PD001867; BPS\_synthase; 1.  
 DR TIGRFAMs: TIGR01072; murA; 1.  
 KM Transferrase; Peptidoglycan synthesis; Cell wall; Cell division;  
 KM Complete proteome.  
 SQ SEQUENCE 430 AA; 45434 MW; 48478C1A34167A25 CRC64;

Query Match 24.1%; Score 49; DB 1; Length 430;  
 Best Local Similarity 39.5%; Pred. No. 19;  
 Matches 15; Conservative 7; Mismatches 14; Indels 2; Gaps 1;

QY 3 LGSRRVFSKCKRLVFSWALVSATMAVTVTLLENTA 40  
 DB 163 LIGARVFPKVS--VGATVLMVAATLANGTTLGNAA 198

RESULT 10  
 RFAE\_HABIN  
 ID "RFAE\_HABIN STANDARD; PRT; 476 AA.  
 AC 005074;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADP-heptose synthase (EC 2.7.-.-).  
 GN RFAE OR WAAB OR H11526.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 OC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxId=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RA MEDLINE=95350630; PubMed=7542800;  
 RA Fieischmann R.D., Bulc C.J., Tomb J.-F., Dougherty B.A., Merrick E.F.,  
 RA Kerlavang A.R., Bulc C.J., Tomb J.-F., Dougherty B.A., Merrick E.F.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodok A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utechtbeck T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Pile L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RL Rd".  
 RL Science 269:496-512(1995).  
 CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.  
 CC -1- SIMILARITY: BELONGS TO THE PRFB FAMILY OF CARBOHYDRATE KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U32828; AAC23172.1; -.  
 DR PIR: C64127; C64127.  
 DR HSP: P27623; 1C0Z.  
 DR TIGR: H11526; -. 1  
 DR InterPro: IPR004821; Cyt tran\_rel.  
 DR InterPro: IPR004820; Cytidylyltransf.  
 DR InterPro: IPR002173; PFKB.  
 DR Pfam: PF01467; CTP transf\_2; 1.  
 DR Pfam: PF00294; PFKB; 1.  
 DR TIGRFAMs: TIGR00125; Cyt tran\_rel; 1.  
 DR PROSITE: PS00583; PFKB\_KINASES\_1; FALSE\_NEG.  
 DR PROSITE: PS00584; PFKB\_KINASES\_2; FALSE\_NEG.  
 KM Lipopolysaccharide biosynthesis; Transferrase; Complete proteome.  
 SQ SEQUENCE 476 AA; 51945 MW; 4F241C08D2C6951E CRC64;

Query Match 24.1%; Score 49; DB 1; Length 476;  
 Best Local Similarity 35.7%; Pred. No. 21;  
 Matches 15; Conservative 4; Mismatches 21; Indels 2; Gaps 1;

QY 4 LGSRRVFSKCKRLVFSWALVS--ATMAVTVTLLENTALAR 43  
 DB 274 LADGRSFESCYLANVAAGIVGKLTSTVTELENAHAR 315

RESULT 11  
 BIR3\_MOUSE  
 ID "BIR3\_MOUSE STANDARD; PRT; 600 AA.  
 AC 008863;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Baculoviral IAP repeat-containing protein 3 (Inhibitor of apoptosis  
 DE protein 1) (MIAP1) (MIAP-1).  
 GN BIRC3 OR BIRC2 OR IAP1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxId=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RA MEDLINE=98110590; PubMed=9441758;  
 RA Liston P., Lefebvre C., Fong W.G., Xuan J.Y., Korneluk R.G.,  
 RA "Genomic characterization of the mouse inhibitor of apoptosis protein  
 RT 1 and 2 genes".  
 RL Genomics 46:495-503(1997).  
 CC -1- FUNCTION: APOPTOTIC SUPPRESSOR. THE BIR MOTIFS REGION INTERACTS  
 CC WITH TNF RECEPTOR ASSOCIATED FACTORS 1 AND 2 (TRAF1 AND TRAF2) TO  
 CC FORM AN HETEROOMERIC COMPLEX, WHICH IS THEN RECRUITED TO THE TUMOR  
 CC NECROSIS FACTOR RECEPTOR 2 (TNFR2) (BY SIMILARITY).  
 CC -1- SUBUNIT: Interacts with SMAC and with PRSS5; these interactions  
 CC inhibit apoptotic suppressor activity (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE IAP FAMILY.  
 CC -1- SIMILARITY: Contains 3 BIR repeats.  
 CC -1- SIMILARITY: Contains 1 CARD domain.  
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U88908; AAC5351.1; -.  
 DR HSP; O13490; 10BH.  
 DR MGD; MG1.197007; Birc2.  
 DR InterPro; IPR001370; BIR.  
 DR InterPro; IPR001315; CARD.  
 DR InterPro; IPR001841; znf\_ring.  
 DR Pfam; PF00653; BIR; 3.  
 DR Pfam; PF00619; CARD; 1.  
 DR Pfam; PF00097; zf-CHC4; 1.  
 DR SMART; SM00238; BIR; 3.  
 DR SMART; SM00114; CARD; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS01282; BIR\_REPEAT\_1; 3.  
 DR PROSITE; PS50143; BIR\_REPEAT\_2; 3.  
 DR PROSITE; PS50209; CARD; 1.  
 DR PROSITE; PS50518; zf\_RING\_1; FALSE\_NEG.  
 DR PROSITE; PS50089; zf\_RING\_2; 1.  
 DR Apoptosis; Zinc-finger; Repeat.  
 KW REPEAT 27 94 BIR 1.  
 FT REPEAT 157 23 BIR 2.  
 FT REPEAT 253 320 BIR 3.  
 FT DOMAIN 436 525 CARD.  
 FT ZN FING 553 588 RING-TYPE.  
 SQ SEQUENCE 600 AA; 67198 MW; AD7F73B6849317D1 CRC64;

Query Match 24.1%; Score 49; DB 1; Length 600;  
 Best Local Similarity 31.6%; Pred. No. 27;  
 Matches 12; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

3 ILGSRVFSKCRVLKFSWALVSATMAVTTLTNTA 40  
 465 LISAPATTEQECNAVKOKPHITQASTLIDTVAKGNIA 502

RESULT 12  
 KG88\_HUMAN  
 ID\_KG88\_HUMAN STANDARD; PRT; 1083 AA.  
 AC O9C0H5;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein KIAA1688.  
 GN KIAA1688.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 RX NCBI\_TaxID=9606;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21082932; PubMed=11214970;  
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XIX.  
 The complete sequences of 100 new cDNA clones from brain which code  
 for large proteins in vitro.";  
 RL DNA Res. 7:347-355(2000).  
 CC -1- SIMILARITY: Contains 1 MYTH4 domain.  
 CC -1- SIMILARITY: Contains 1 Rho-GAP domain.  
 CC -1- SIMILARITY: Contains 2 WW domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB051475; BAB21779.1; ALT\_INIT.  
 DR HSP; O07960; TRGP.  
 DR InterPro; IPR000857; MYTH4.  
 DR InterPro; IPR001198; RhogAP.  
 DR InterPro; IPR001202; WW\_Rp5\_WMP.  
 DR Pfam; PF00784; MYTH4; 1.  
 DR Pfam; PF00620; RhogAP; 1.  
 DR Pfam; PF00397; WW; 2.  
 DR SMART; SM00139; MYTH4; 1.  
 DR SMART; SM00324; RhogAP; 1.  
 DR SMART; SM00456; WW; 2.  
 DR PROSITE; PS50238; RhogAP; 1.  
 DR PROSITE; PS01159; WW\_DOMAIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50020; WW\_DOMAIN\_2; 1.  
 KW Repeat.  
 FT DOMAIN 25 58 WW 1.  
 FT DOMAIN 63 97 WW 2.  
 FT DOMAIN 761 879 MYTH4.  
 FT DOMAIN 890 1078 RHO-GAP.  
 FT DOMAIN 234 367 PRO-RICH.  
 SQ SEQUENCE 1083 AA; 121285 MW; 782B213B1DCB4072 CRC64;

Query Match 24.1%; Score 49; DB 1; Length 1083;  
 Best Local Similarity 35.7%; Pred. No. 49;  
 Matches 10; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

14 CRLVFSWALVSATMAVTTLTNTAL 41  
 1012 CYLIRFLQVFPQNPANVANTKMDVSNLNM 1039

RESULT 13  
 KR1\_YEAST  
 ID\_KR1\_YEAST STANDARD; PRT; 928 AA.  
 AC P36003;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable serine/threonine-protein kinase YKL171W (EC 2.7.1.-).  
 GN YKL171W OR YKL635.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 RX NCBI\_TaxID=4932;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RX MEDLINE=94378719; PubMed=8091858;  
 RA Vandenbol M., Bolle P.-A., Dion C., Portetelle D., Hilger F.;  
 RT "Sequencing and analysis of a 20.5 kb DNA segment located on the left  
 arm of yeast chromosome XI.";  
 RL Yeast 10:S25-S33(1994).  
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE: PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KM Hypothetical protein; Transferase; Serine/threonine-protein kinase;  
 KM ATP-binding.  
 FT DOMAIN 449 912 PROTEIN KINASE.  
 FT NP BIND 455 463 ATP (BY SIMILARITY).  
 FT BINDING 478 478 ATP (BY SIMILARITY).  
 FT ACT SITE 580 580 POTENTIAL.  
 SQ SEQUENCE 928 AA; 103956 MW; 00EB0353EC5999A CRC64;

Query Match 23.9%; Score 48.5; DB 1; Length 928;  
 Best Local Similarity 31.1%; Pred. No. 49;  
 Matches 14; Conservative 9; Mismatches 15; Indels 7; Gaps 1;

QY 6 SRVPSKRCVAVKFSNVAVSA-----TMAVTTVLTENTLALR 43  
 DB 546 SKIPFAERCRLTFTLSLQLSLAKVMSKTIYHGDIKLENCILQK 590

## RESULT 14

KG88 RAT STANDARD; PRT; 101 AA.  
 ID KG88 RAT  
 AC P18850;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Protein KIAA1688 homolog (Preoptic regulatory factor-2) (PDRF-2) (Fragment).  
 DE Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=91155954; PubMed=2293025;  
 RA Nowak F.V.;  
 RT "Cloning of two hypothalamic cDNAs encoding tissue-specific transcripts in the preoptic area and testis."  
 RL Mol. Endocrinol. 4:1205-1210(1990).  
 CC -1- TISSUE SPECIFICITY: Preoptic area and testis.  
 CC -1- SIMILARITY: Contains 1 Rho-GAP domain.  
 CC -1- CAUTION: Was originally (Ref.1) thought to be a precursor for a secreted GHRH-like peptide.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to frameshifts in positions 13 and 23.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL: X53232; CA37324.1; ALT\_FRAME.  
 DR InterPro: IPR000198; RhoGAP.  
 DR PROSITE: PS50238; RHO GAP; 1.  
 FT NON TER 1 1  
 FT DOMAIN <1 96 RHO-GAP.  
 SQ SEQUENCE 101 AA; 11572 MW; F216805966BE143E CRC64;

Query Match 23.6%; Score 48; DB 1; Length 101;  
 Best Local Similarity 32.1%; Pred. No. 5.9;  
 Matches 9; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

QY 14 CRVKSVMALVSAATMAVTTVLTENTLAL 41  
 DB 30 CYLIRFQVFPVOPANVAITMDVSNILAM 57

## RESULT 15

TRT4 ECOLI STANDARD; PRT; 243 AA.  
 ID TRT4 ECOLI  
 AC P15177;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Trt4 complement resistance protein precursor.  
 GN Trt4.  
 OS Escherichia coli.  
 OC plasmid ColB-K38.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RA O'Connor D.;  
 RL Submitted (MAR-1989) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: RESPONSIBLE FOR PREVENTING UNPRODUCTIVE CONJUGATION BETWEEN BACTERIA CARRYING LIKE PLASMIDS.  
 CC -1- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X14566; CA32704.1; -  
 DR PIR: S07997; S07997.  
 DR PROSITE: PS00013; PROKAR\_LIPOPROTEIN; 1.  
 KM Plasmid; Conjugation; Signal; Lipoprotein; Outer membrane.  
 FT SIGNAL 1 20  
 FT CHAIN 21 243  
 FT LIPID 21 21 N-ACYL DIGLYCERIDE.  
 SQ SEQUENCE 243 AA; 25873 MW; B8305B10FDB1239 CRC64;

Query Match 23.6%; Score 48; DB 1; Length 243;  
 Best Local Similarity 42.3%; Pred. No. 15;  
 Matches 11; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 18 KFSWVAIVSATMAVTTVLTENTLALR 43  
 DB 5 KLMWVAIVSTLALSGCGAMSTAIKK 30

Search completed: January 5, 2004, 18:39:08  
 Job time : 8.64272 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 16:40:38 ; Search time 9.9779 Seconds  
(without alignments)  
1112.084 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203

Sequence: 1 MNLGSRVRSKCKRLVKFS.....LVSATMAVTTVLENTALAR 43

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mbc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	203	100.0	271	16 Q54969	Q54969 streptococc
2	203	100.0	271	16 Q9ACN6	Q9ACN6 streptococc
3	59	29.1	458	16 Q9X227	Q9X227 thermotoga
4	57	28.1	1595	4 Q8IVG6	Q8IVG6 homo sapien
5	56	27.6	268	8 Q8G9E0	Q8G9E0 ascapus tr
6	56	27.6	276	8 Q8G9E3	Q8G9E3 ascapus tr
7	55	27.1	366	17 Q29257	Q29257 archaeoglob
8	55	27.1	576	16 Q9RU44	Q9RU44 deinococcus
9	54	26.6	403	16 Q8EWJ7	Q8EWJ7 mycoplasma
10	53	26.1	209	8 Q8G9D9	Q8G9D9 ascapus tr
11	53	26.1	271	8 Q8G9E2	Q8G9E2 ascapus tr
12	53	26.1	276	8 Q8G9E1	Q8G9E1 ascapus tr
13	53	26.1	276	8 Q8G1E5	Q8G1E5 ascapus tr
14	53	26.1	276	8 Q8G1E8	Q8G1E8 ascapus tr
15	53	26.1	310	16 Q984X2	Q984X2 rhizobium l
16	53	26.1	361	5 Q9U998	Q9U998 anopheles f

17	53	26.1	628	17 Q8PTP4	Q8PTP4 methanosarc
18	53	26.1	827	16 Q8PE57	Q8PE57 xanthomonas
19	53	26.1	1051	5 Q9VA99	Q9VA99 drosophila
20	53	26.1	1109	5 Q8IMI9	Q8IMI9 drosophila
21	52.5	25.9	908	5 O01598	O01598 caenorhabdi
22	52	25.6	215	10 Q41061	Q41061 pisum sativ
23	52	25.6	341	8 Q8HC29	Q8HC29 menura nova
24	52	25.6	467	16 Q8UC28	Q8UC28 agrobacteri
25	51.5	25.4	178	2 O85185	O85185 escherichia
26	51.5	25.4	206	16 Q55620	Q55620 synechocyst
27	51.5	25.4	221	16 P74602	P74602 synechocyst
28	51.5	25.4	728	16 O8X606	O8X606 escherichia
29	51	25.1	481	2 Q9X4U5	Q9X4U5 streptococc
30	51	25.1	658	10 O48788	O48788 arabidopsis
31	51	25.1	658	10 Q949V3	Q949V3 arabidopsis
32	51	25.1	730	5 Q9VNL6	Q9VNL6 drosophila
33	51	25.1	827	16 Q8PCF8	Q8PCF8 xanthomonas
34	51	25.1	838	16 O84625	O84625 chlamydia t
35	51	25.1	1042	16 Q97G63	Q97G63 clostridium
36	50.5	24.9	269	16 Q8XEQ5	Q8XEQ5 salmonella
37	50.5	24.9	458	2 Q8GDU5	Q8GDU5 heliobacill
38	50.5	24.9	598	5 Q21035	Q21035 caenorhabdi
39	50	24.6	126	16 Q9ADN1	Q9ADN1 streptomyce
40	50	24.6	243	2 Q9EUJ2	Q9EUJ2 salmonella
41	50	24.6	271	12 Q69374	Q69374 gallid herp
42	50	24.6	473	16 Q8F399	Q8F399 leptospira
43	50	24.6	537	12 Q91HJ1	Q91HJ1 gallid herp
44	50	24.6	888	16 Q98PQ2	Q98PQ2 mycoplasma
45	50	24.6	891	16 Q8EBE8	Q8EBE8 shewanella

## ALIGNMENTS

### RESULT 1

Q54969	ID	Q54969	PRELIMINARY;	PRT;	271 AA.
AC	Q54969;				
DT	01-NOV-1996	(TrEMBLrel. 01, Created)			
DT	01-NOV-1996	(TrEMBLrel. 01, Last sequence update)			
DT	01-OCT-2002	(TrEMBLrel. 22, Last annotation update)			
DE	Mitogenic factor precursor (Mitogenic factor, 25K).				
OS	MF OR SPY2043 OR SPYMI8_2104.				
GN	Streptococcus pyogenes, and				
OS	Streptococcus pyogenes (serotype M18).				
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
OC	Streptococcus.				
OX	NCBI_TaxID=1314, 186103;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RC	STRAIN=NY-5;				
RX	MEDLINE=94009636; PubMed=8405402;				
RA	Iwasaki M., Igarashi H., Hinuma Y., Yutsudo T.;				
RT	"Cloning, characterization and overexpression of a Streptococcus pyogenes gene encoding a new type of mitogenic factor.";				
RL	FEBS Lett. 331:187-192(1993).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SV;				
RA	Hong K.;				
RT	"A novel cloning method used arbitrarily primed PCR.";				
RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=SF370 / ATCC 700294 / Serotype M1;				
RX	MEDLINE=21192694; PubMed=11296296;				
RA	Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,				
RT	Primeaux C., Szate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,				
RA	Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,				
RA	Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;				
RT	"Complete genome sequence of an M1 strain of Streptococcus pyogenes.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).				
RN	[4]				

```
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; D13428; BAA02693.1; --
DR EMBL; AB030578; BAB16025.1; --
DR EMBL; AE006625; AAK34710.1; --
DR EMBL; AE010113; AAL98563.1; --
KW Nuclease; Signal; Complete proteome.
FT SIGNAL 1 43 POTENTIAL.
FT CHAIN 44 271 MITOGENIC FACTOR.
SQ SEQUENCE 271 AA; 30062 MW; 8FD89DA38A124352 CRC64;

Query Match 100.0%; Score 203; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 8.3e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSWALVSATWATVTTVLTALAR 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATWATVTTVLTALAR 43

RESULT 2
Q9ACN6 PRELIMINARY; PRT; 271 AA.
AC Q9ACN6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Streptodornase B precursor (Mitogenic factor 25K).
GN DNASEB OR MF OR SPY3 1745.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 198466;
RN 11 PRELIMINARY; PRT; 271 AA.
RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN=C203S;
RX MEDLINE=21233096; PubMed=11335140;
RA Gerlach D., Schmidt K.H., Fleischer B.;
RA "Basic streptococcal superantigens (SPEX/SMEZ or SPEC) are responsible
RT for the mitogenic activity of the so called mitogenic factor MF.";
RL FEMS Immunol. Med. Microbiol. 30:209-216 (2001).
RN 12

RP SEQUENCE FROM N.A.
RC SPECIES=S.pyogenes; STRAIN-MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormack J.K., Leung D.Y.W.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; AJ295272; CAC35734.1; --
DR EMBL; AE014170; AAM80352.1; --
KW Signal; Complete proteome.
FT SIGNAL 1 42 POTENTIAL.
FT CHAIN 43 271 STREPTODORNASE B.
SQ SEQUENCE 271 AA; 30090 MW; 8FD8849E3CD01352 CRC64;

Query Match 100.0%; Score 203; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 8.3e-21;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MNLGSRVFSKCKRLVKFSWALVSATWATVTTVLTALAR 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATWATVTTVLTALAR 43

RESULT 3
Q9X227 PRELIMINARY; PRT; 458 AA.
AC Q9X227;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein TM1701.
GN TM1701.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:323-329 (1999).
DR EMBL; AE001810; AAD36768.1; --
DR TIGR; TM1701.
DR InterPro; IPR002528; Mate.
DR InterPro; IPR001014; Ribosomal_L23.
DR Pfam; PF01554; Mate; 2.
DR TIGRFAMs; TIGR00797; mate; 1.
DR PROSITE; PS00050; RIBOSOMAL_L23; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 458 AA; 49857 MW; 7C7E7D07F7DD4543 CRC64;

Query Match 29.1%; Score 59; DB 16; Length 458;
Best Local Similarity 44.4%; Pred. No. 5;
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 9 VFSSKCKRLVKFSWALVSATWATVTTVT 35
DB 276 VFTSAMRVINFGTVPLIGMAMAVTSVT 302

RESULT 4
Q8IVG6 PRELIMINARY; PRT; 1595 AA.
AC Q8IVG6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA2004 (Fragment).
GN KIAA2004.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095925; BAC23101.1; --
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1595 AA; 184969 MW; 1DC92BE68887B662 CRC64;
```

	Query Match	28.1%	Score 57	DB 4	Length 1595
	Best Local Similarity	34.3%	Pred. No. 33		
	Matches 12	Conservative 13	Mismatches 8	Indels 2	Gaps 1
Qy	2	NLGSRRRVFKCKRLVKFMSVALVSATMAVTVTL	36		
Db	909	NILKGNQIFKEAKL--FSFLALNSVVPDTISL	941		

```

RESULT 5
QG9GEO PRELIMINARY; PRT; 268 AA.
AC QSG9EO;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 2 (Fragment).
GN ND2.
OS Ascaphus truei (tailed frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Ascaphidae; Ascaphus.
OX NCBI_TaxID=8439;
[1]
RN SEQUENCE FROM N.A.
RP Nielson M., Lohman K., Sullivan J.;
RA "Phylogeography of the Tailed Frog (Ascaphus truei): Implications for
RT the Biogeography of the Pacific Northwest.";
RL Evolution 0:0-0(2001).
RR EMBL; AF277365; AAC26386.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON_TER 268
SQ SEQUENCE 268 AA; 229449 MW; 3A81119942D5AE05 CRC64;

```

```

Query Match      27.6%; Score 56; DB 8; Length 268;
Best Local Similarity 32.4%; Pred.No. 7.9;
Matches 11; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy      9 VFSSKCKLRFPSMVALVSATMAVTTVLTENTALA 42
      :  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      195 ILAPKMLLNLSMYLLLTSTMFMTLTTTKIS 238

```

```

RESULT 6
QG9G9E3
ID QG9E3 PRELIMINARY; PRT; 276 AA.
AC QG9E3;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE NADH dehydrogenase subunit 2 (Fragment).
DN ND2.
GN Ascapush truei (tailed frog).
OS Mitochondrion.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Ascapidae; Ascapush.
OX NCBI_TaxID=8439;
[1]
SEQUENCE FROM N.A.
RP Nielson M., Lohman K., Sullivan J.;
RT "Phylogeography of the Tailed Frog (Ascapush truei): Implications for
RT the Biogeography of the Pacific Northwest.";
RL Evolution 0:0-0(2001).
RR ENBL; AF277355; AAG26376.1; -.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; oxidored_q1; 1.
KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT NON_TER 276
SQ SEQUENCE 276 AA; 30360 MW; 4B4F3E3EBCE9310C CRC64;

```

Query Match 27.6%; Score 56; DB 8; Length 276;

Best Local Similarity 32.4%; Pred. No. 8.2;  
Matches 11; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy 9 VFSSKCLVFSMVALVSATMAVTITLTALTA 42  
: : | - | : : : : | : | :  
Db 195 ILAPKMLNLSMYLLSTMFMTITLTTTKIS 228

```

RESULT 7
029257 ID 029257 PRELIMINARY; PRT; 366 AA.
AC AC 029257;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE ABC transporter, ATP-binding protein, putative.
GN AFI005.
OS Archaeoglobus fulgidus.
OC Archaea, Euryarchaeota; Archaeoglobi; Archaeoglobates;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.T., McNeil L.K., Badger J.H., Glodek A., Zhou
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artchak P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus." ;
RL Nature 390:364-370(1997).
RL DR EMBL; AE001034; AAB90233.1; -.
DR TIGR; AF1005; -.
DR InterPro; IPR000412; ABC_transp2.
DR PRINTS; PR00164; ABC2TRNSPORT.
DR Hypothetical protein; ATP-binding; Complete proteome.
KW SO SEQUENCE 366 AA; 39790 MW; 46c9b836d02a9c8c CR664;

```

```
Query Match      27.1%; Score 55; DB 17; Length 366;
Best Local Similarity 29.7%; Pred. No. 15;
Matches 11; Conservative 10; Mismatches 16; Indels 0; Gaps 0;
```

```

RESULT 8
Q9RU44
ID          Q9RU44          PRELIMINARY;          PRT;          576 AA.
AC          Q9RU44;
DT          01-MAY-2000 (TTEMBLrel. 13, Created)
DT          01-MAY-2000 (TTEMBLrel. 13, Last sequence update)
DT          01-MAR-2003 (TTEMBLrel. 23, Last annotation update)
DE          F7SE protein.
GN          DR1550.
OS          Deinococcus radiodurans.
OC          Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC          Deinococcaceae; Deinococcus.
ON          NCBI_TaxID=1299;
RX          [1]
RA          SEQUENCE FROM N.A.
RC          STRAIN=R1;
RX          MEDLINE=20036896; PubMed=10567266;
RA          White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peter

```

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans R1.";  
 RL Science 286:1571-1577 (1999).  
 DR EMBL; AB001998; AAF11110.1; --  
 DR TIGR; DR1550; --  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR003439; ABC transporter.  
 DR InterPro; IPR003838; DUF214.  
 DR InterPro; IPR005286; IISF.  
 DR Pfam; PF00005; ABC\_tran; 1.  
 DR Pfam; PF02687; DUF214; 1.  
 DR ProDom; PD000006; ABC\_transporter; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR TIGRFAMs; TIGR00960; 3a0501802; 1.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 576 AA; 63413 MW; 6167B2181D583DDC CRC64;

Query Match 27.1%; Score 55; DB 16; Length 576;  
 Best Local Similarity 35.7%; Pred. No. 23;  
 Matches 15; Conservative 8; Mismatches 19; Indels 0; Gaps 0;  
 QY 1 MNLGSRVFKKRLVKFMSVALVSATMAVTTVLENTALA 42  
 Db 485 MRLGATRSFIRPHLIEGLLVGVAALVATLAWGQVLA 526

RESULT 9  
 Q8EWJ7 PRELIMINARY; PRT; 403 AA.  
 AC Q8EWJ7;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical protein.  
 GN MYPE2060.  
 OS Mycoplasma penetrans.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=28227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HF-2;  
 RX MEDLINE=22354719; PubMed=12466555;  
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,  
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;  
 RT "The complete genomic sequence of Mycoplasma penetrans, an  
 RT intracellular bacterial pathogen in humans.";  
 RL Nucleic Acids Res. 30:5293-5300 (2002).  
 DR EMBL; AP004170; BAC43997.1; --  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 403 AA; 43742 MW; E65A78BA427BF54E CRC64;

Query Match 26.6%; Score 54; DB 16; Length 403;  
 Best Local Similarity 52.2%; Pred. No. 23;  
 Matches 12; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
 QY 18 KFSMVLSATMAVTTVLENTA 40  
 Db 8 KFIKFLASSTVLTITLSSCA 30

RESULT 10  
 Q9G9D9 PRELIMINARY; PRT; 209 AA.  
 AC Q9G9D9;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE NADH dehydrogenase subunit 2 (Fragment).  
 GN ND2.  
 OS Ascaphus truei (tailed frog).  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Ascaphidae; Ascaphus.  
 OX NCBI\_TaxID=8439;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nielsen M., Lohman K., Sullivan J.;  
 RT "Phylogeography of the Tailed Frog (Ascaphus truei): Implications for  
 RT the Biogeography of the Pacific Northwest.";  
 RL Evolution 0:0-0(2001).  
 DR EMBL; AF277367; AAG26380.1; --  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR Pfam; PF00361; oxidored\_q1; 1.  
 DR KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.  
 FT NON TER 1 209  
 FT NON TER 209  
 SQ SEQUENCE 209 AA; 22921 MW; F5A6039406541453 CRC64;

Query Match 26.1%; Score 53; DB 8; Length 209;  
 Best Local Similarity 29.4%; Pred. No. 17;  
 Matches 10; Conservative 11; Mismatches 13; Indels 0; Gaps 0;  
 QY 9 VFSKCKRLVKFMSVALVSATMAVTTVLENTALA 42  
 Db 148 ILAPKLVLNLSMYLLLTSTMTFLTMITLTITTKIS 181

RESULT 11  
 Q9G9E2 PRELIMINARY; PRT; 271 AA.  
 AC Q9G9E2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE NADH dehydrogenase subunit 2 (Fragment).  
 GN ND2.  
 OS Ascaphus truei (tailed frog).  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Archeobatrachia; Ascaphidae; Ascaphus.  
 OX NCBI\_TaxID=8439;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nielsen M., Lohman K., Sullivan J.;  
 RT "Phylogeography of the Tailed Frog (Ascaphus truei): Implications for  
 RT the Biogeography of the Pacific Northwest.";  
 RL Evolution 0:0-0(2001).  
 DR EMBL; AF277359; AAG26380.1; --  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR Pfam; PF00361; oxidored\_q1; 1.  
 DR KW NAD; Oxidoreductase; Ubiquinone; Mitochondrion.  
 FT NON TER 271  
 SQ SEQUENCE 271 AA; 29779 MW; C5DFFE3A81046CA2 CRC64;

Query Match 26.1%; Score 53; DB 8; Length 271;  
 Best Local Similarity 29.4%; Pred. No. 21;  
 Matches 10; Conservative 11; Mismatches 13; Indels 0; Gaps 0;  
 QY 9 VFSKCKRLVKFMSVALVSATMAVTTVLENTALA 42  
 Db 195 ILAPKLVLNLSMYLLLTSTMTFLTMITLTITTKIS 228

RESULT 12  
 Q9G9E1 PRELIMINARY; PRT; 276 AA.  
 AC Q9G9E1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	NADH dehydrogenase subunit 2 (Fragment).
GN	ND2.
OS	Ascaphus truei (tailed frog).
OC	Mitochondrion.
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Archeobatrachia; Ascaphidae; Ascaphus.
OX	NCBI_TaxID=8439;
[1]	
RN	SEQUENCE FROM N.A.
RA	Nielson M., Lohman K., Sullivan J.;
FT	"Phylogeography of the Tailed Frog (Ascaphus truei): Implications for
RT	the Biogeography of the Pacific Northwest.";
RL	Evolution 0:0-0(2001).
DR	EMBL; AF277364; AAG26385.1; -.
DR	InterPro; IPR001750; Oxidored_q1.
DR	Pfam; PF00361; oxidored_q1_1.
KW	NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT	NON TER 276 276
SQ	SEQUENCE 276 AA; 30314 MW; B6F485CDF0068604 CRC64;
Query Match	26.1%; Score 53; DB 8; Length 276;
Best Local Similarity	29.4%; Pred.No. 22;
Matches 10; Conservative 11; Mismatches 13; Indels 0; Gaps 0;	
Oy	9 VFSSKCRLVKFMSVALVSATMAVTTVTLENTALA 42
:	: : :
DB	195 ILAPKLVLNLNSMYLLTSTMFMTLTITTKIS 228
RESULT 13	
Q9G1C5	PRELIMINARY; PRT; 276 AA.
ID	Q9G1C5
AC	Q9G1CS;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	NADH dehydrogenase subunit 2 (Fragment).
GN	ND2.
OS	Ascaphus truei (tailed frog).
OC	Mitochondrion.
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Archeobatrachia; Ascaphidae; Ascaphus.
OX	NCBI_TaxID=8439;
[1]	
RN	SEQUENCE FROM N.A.
RA	Nielson M., Lohman K., Sullivan J.;
FT	"Phylogeography of the Tailed Frog (Ascaphus truei): Implications for
RT	the Biogeography of the Pacific Northwest.";
RL	Evolution 0:0-0(2001).
DR	EMBL; AF277366; AAG26387.1; -.
DR	EMBL; AF277356; AAG26377.1; -.
DR	InterPro; IPR001750; Oxidored_q1.
DR	Pfam; PF00361; oxidored_q1_1.
KW	NAD; Oxidoreductase; Ubiquinone; Mitochondrion.
FT	NON TER 276 276
SQ	SEQUENCE 276 AA; 30346 MW; A29167DF46FEBA4B CRC64;
Query Match	26.1%; Score 53; DB 8; Length 276;
Best Local Similarity	29.4%; Pred.No. 22;
Matches 10; Conservative 11; Mismatches 13; Indels 0; Gaps 0;	
Oy	9 VFSSKCRLVKFMSVALVSATMAVTTVTLENTALA 42
:	: : :
DB	195 ILAPKLVLNLNSMYLLTSTMFMTLTITTKIS 228
RESULT 14	
Q9GIY8	PRELIMINARY; PRT; 276 AA.
ID	Q9GIY8
AC	Q9GIY8;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

Db 132 RLDRLQIAA VVLATVAVTVLTIEGKL 158

Search completed: January 5, 2004, 18:41:29  
Job time : 16.9779 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:53:50 ; Search time 3989 Seconds  
(without alignments)  
11106.825 Million cell updates/sec

Title: US-08-482-785-7

Perfect score: 1083

Sequence: 1 GACACGCCCTCTCTTTTCT.....TGTGCAAAAGCAAAAGC 1083

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_on.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1083	100.0	1083	6	A49201 Sequence 7
2	1083	100.0	1083	6	AR219224 Sequence
3	1046.8	96.7	53291	1	AE014170 Streptoco
4	1046.8	96.7	323825	1	AP005146 Streptoco
5	1044.4	96.4	12808	1	AE006625 Streptoco
6	1042.8	96.3	2099	1	STRMP
7	1037.2	95.8	10424	1	AE010113 Streptoco
8	1030.8	95.2	4334	1	AB030578 Streptoco
9	961.6	88.8	1021	6	E08254 DNA encodin
10	922.6	85.2	940	6	A49205 Sequence 11
11	922.6	85.2	940	6	AR219226 Sequence
12	853.8	78.8	937	6	A49208 Sequence 14
13	853.8	78.8	937	6	AR219229 Sequence
14	816	75.3	816	1	SPY295272 Streptoco
15	276.8	25.6	50081	1	AE014158 Streptoco
16	276.8	25.6	310950	1	AP005143 Streptoco
17	200	18.5	200	6	A49204 Sequence 10
18	200	18.5	200	6	AR219225 Sequence
19	128	11.8	14663	1	AE010084 Streptoco
20	127.4	11.8	2147	1	SPSDAGENE
21	106.6	9.8	182	6	A49206 Sequence 12
22	106.6	9.8	182	6	AR219227 Sequence
23	88.4	8.2	1023	1	AF410852 Streptoco
24	88.4	8.2	1113	1	SESDC
25	88.4	8.2	310850	1	AP005142 Streptoco
26	86.8	8.0	50820	1	AE014163 Streptoco
27	72	6.6	10029	1	AE010112 Streptoco
28	70.8	6.5	324050	1	AL935253 Lactobaci
29	47.2	4.4	1141	6	AX083744 Sequence
30	46.2	4.3	6106	6	AX251183 Sequence
31	46.2	4.3	6106	6	AX346374 Sequence
32	46.2	4.3	6106	6	AX348418 Sequence
33	45	4.2	7218	6	I66494 Sequence 14
34	45	4.2	349980	6	AX344564 Sequence
35	44.4	4.1	3334	3	AY254474 Dictyoste
36	44.4	4.1	62927	8	AB018119 Arabidops
37	44	4.1	2406	6	AX141367 Sequence
38	44	4.1	2942	1	AF270390 Staphyloc
39	44	4.1	2942	6	AX145708 Sequence
40	44	4.1	2976	1	AF270160 Staphyloc
41	44	4.1	2976	6	AX145478 Sequence
42	44	4.1	3350	1	AF269617 Staphyloc
43	44	4.1	3350	6	AX144937 Sequence
44	44	4.1	6013	6	AX251964 Sequence
45	44	4.1	6013	6	AX344356 Sequence

# ALIGNMENTS

## RESULT 1

A49201	A49201	1083 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	Sequence 7 from Patent WO9606174.				
DEFINITION	A49201				
ACCESSION	A49201				
VERSION	A49201.1	GI:2302733			
KEYWORDS	Streptococcus pyogenes				
SOURCE	Streptococcus pyogenes				
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;				
REFERENCE	1 (bases 1 to 1083)				
AUTHORS	Adams,C.W., Pang,P.P. and Belei,M.C.				
TITLE	RECOMBINANT DNase B DERIVED FROM STREPTOCOCCUS PYOGENES				
JOURNAL	Patent: WO 9606174-A 7 29-FEB-1996;				

[illegible]



```
Db 301 GCAAGTACTAAACGAACATTAGCTTGACATTCATGACAGTCCCTNACTATTACAAA 360
Qy 361 CTTTAGGTACTAGTCAGATTACTCAGACATCTTTCTTCTTAAAGCAGGAGATTTCTCTATA 420
Db 361 CTTTAGGTACTAGTCAGATTACTCAGACATCTTTCTTCTTAAAGCAGGAGATTTCTCTATA 420
Qy 421 GCAATATAGATGAGTTAGGAGGAGCGCTACTGCTAGAGGTACATTGACTTATGCCAATG 480
Db 421 GCAATATAGATGAGTTAGGAGGAGCGCTACTGCTAGAGGTACATTGACTTATGCCAATG 480
Qy 481 TTGAAGGTAGTACGCTGTAGACAAATCTTTCCGTAAAAATCAAAACCCCGCAGATGGA 540
Db 481 TTGAAGGTAGTACGCTGTAGACAAATCTTTCCGTAAAAATCAAAACCCCGCAGATGGA 540
Qy 541 CTGGAACCCCTTAATCATGATCAAAATATAAAATGAAATGGTTAAATGGTCTATCTTATGTCG 600
Db 541 CTGGAACCCCTTAATCATGATCAAAATATAAAATGAAATGGTTAAATGGTCTATCTTATGTCG 600
Qy 601 GAGATTTCTGAATAGAGTATCATCTCATGTCAGATAGTCTCGGTGGAGATGCACCTCAGAG 660
Db 601 GAGATTTCTGAATAGAGTATCATCTCATGTCAGATAGTCTCGGTGGAGATGCACCTCAGAG 660
Qy 661 TCAATGCCGTTTACAGGAACAGTACCCAAAATGTAGGAGTCTGTGACCAAAAAGCGCGCA 720
Db 661 TCAATGCCGTTTACAGGAACAGTACCCAAAATGTAGGAGTCTGTGACCAAAAAGCGCGCA 720
Qy 721 TGCCTATACCGAACAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780
Db 721 TGCCTATACCGAACAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780
Qy 781 ATTATGAAGTCGCTCCCAATCTACAAACGAGACGATGATTTCAAAGAGCTGTCTGGTAT 840
Db 781 ATTATGAAGTCGCTCCCAATCTACAAACGAGACGATGATTTCAAAGAGCTGTCTGGTAT 840
Qy 841 CAATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 900
Db 841 CAATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 900
Qy 901 GCTACACCATTAATACCATACCGTACACCTACTCAAAAATAATACCAAAAGGCTAGAC 960
Db 901 GCTACACCATTAATACCATACCGTACACCTACTCAAAAATAATACCAAAAGGCTAGAC 960
Qy 961 CTCTGCTCCTAGCGCTAGCTTTTATACATCAAAAAGCAATGACTATAGAAAGTAAAAA 1020
Db 961 CTCTGCTCCTAGCGCTAGCTTTTATACATCAAAAAGCAATGACTATAGAAAGTAAAAA 1020
Qy 1021 TACTAGAAAAGCAATGATTCCTGCTCATGCTTTTATGATTTGTCGCAAAAGCAAAAA 1080
Db 1021 TACTAGAAAAGCAATGATTCCTGCTCATGCTTTTATGATTTGTCGCAAAAGCAAAAA 1080
Qy 1081 AGC 1083
Db 1081 AGC 1083
```

```
RESULT 3
AE014170/c
LOCUS AE014170 53291 bp DNA linear BCT 19-JUL-2002
DEFINITION Streptococcus pyogenes MGAS315, section 35 of 37 of the complete
Genome.
ACCESSION AE014170 AE014074
VERSION AE014170.1 GI:21905475
SOURCE Streptococcus pyogenes MGAS315
ORGANISM Streptococcus pyogenes MGAS315
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 53291)
Beres, S.B., Sylva, G.L., Barbican, K.D., Lei, B., Hoff, J.S.,
Mammarella, N.D., Liu, M.-Y., Smoot, J.C., Porcella, S.F.,
Parkins, L.D., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M. and
Musser, J.M.
```

## TITLE

JOURNAL  
REFERENCE  
AUTHORS

TITLE  
JOURNAL

FEATURES  
source

gene

CDS

gene

CDS

Genome sequence of a serotype M3 strain of group A Streptococcus:  
Phage-encoded toxins, the high-virulence phenotype, and clone  
emergence  
Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10078-10083 (2002)  
2 (bases 1 to 53291)  
Beres, S.B., Sylva, G.L., Barbican, K.D., Lei, B., Hoff, J.S.,  
Mammarella, N.D., Liu, M.-Y., Smoot, J.C., Porcella, S.F.,  
Parkins, L.D., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M. and  
Musser, J.M.  
Direct Submission  
Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis,  
Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St.,  
Hamilton, MT 59840, USA  
Location/Qualifiers  
1. 53291  
/organism="Streptococcus pyogenes MGAS315"  
/mol\_type="genomic DNA"  
/strain="MGAS315"  
/serotype="M3"  
/db\_xref="taxon:198466"  
/notes="group: A"  
complement (162.3671)  
/genes="scpA"  
/notes="synonym: SpvM3.1726"  
complement (162.3671)  
/genes="scpA"  
/notes="best blastp hit: sp|P15926|SCA1\_STRPY CSA peptidase  
precursor (SCP)"  
/codon\_start=1  
/transl\_table=11  
/product="CSA peptidase precursor"  
/protein\_id="AA080333.1"  
/db\_xref="GI:21905476"  
/translation="MRKKOKLPDKLAALMSTSLINAQSDIKANTVTDTPATEQA  
VEVQQTAVSEEPSSSSKEDNPOTPDAAETVADKANDLAPOAPAKTADIPATSE  
TIRDLDPDSHVKTLOEKACGACTVAVIDAGDKHAEARLTDKSKARQSKEDLEK  
AKDGGITGGEWVDKAYHYDYSKDKTAVDQEHGTHVSGILSNAPSEKTEPRLE  
GAMPEAQLLMRVEIVNGLADYARNAQAIIRDANLGAIVINSGFNAALAYANLPDE  
TKAFDYAKSGVSI VTSAGNDSFGGCTRLPLADHPDYGVTGTPAAADSTLVASYSD  
IKGLDIDFKDKIANAKKAGAVGLIYDNDKGPPIELPNVDQMPAAFIKRDGILLKD  
PERQETATKTDADQDKEMVLSYRPNPKAYDYVANRGTKEDDFKDKGKIAL  
NSKTTITFNATPKVLTASGKLSRSGSLGTADGNIKPDIAAPGQDILSSVANNKYA  
KLSGTSAPLAVAGIMGLIQOYETQPDMPSPERLDLAKKLVSSALATYDEDEKAY  
FSRQCGAGADAKASAAATVYTDKNTSSKVLANNVSDKFEVTVTHNKSQPEL  
YYQATVQTKVDGKHPALAPKALYETSWKITIPANSKQVTPVIDASRFSKDLAQM  
KNGYFLGFRPKQDPTKELMSIPYIGRDFGNLSALEKPIYDSKDGSSYYHEANS  
DKAQDGDGLQFYALKNNFTALTESNPWTIIKAVKEGVENIEDIESSEITETIPAG  
TFAKQDSDSHYIHRHANGKPYAAISPNGDGRDYVQPGTFLRNKALVAEVLDEKG  
DVVMTSEVTEQVYVNNNDLSTGLSTRFEKTRWDGDKDGKVVANGTYTYRVRTPI  
SSGAKQHTDFDIVDNTTPEAATSAATSAEDRRLTLASKPSTQPYRRIATYMD  
EDLPTTEYISPNEDGTFLLPEEAETMEGATVPLKMSDFTYVEDMAGNIYTYPTVNL  
EGHSNKPEODSDQVDPKTPPEODGSGQAPDKKPEAKPEODGSGQAPDKKPEKTP  
EKDSGQTPGKTPKQPSRTLEKRSKRALATKASARDQLPTNDKDTNRLHLLLV  
MTTFFGLVAHFRTKQKETKK"  
complement (4041.5786)  
/genes="emm3"  
/notes="synonym: SpvM3.1727"  
complement (4041.5786)  
/genes="emm3"  
/notes="best blastp hit: emb|CAA56449.1| (X80168) M3  
protein [Streptococcus pyogenes]"  
/codon\_start=1  
/transl\_table=11  
/product="antiphagocytic M protein, type 3"  
/protein\_id="AA080334.1"  
/db\_xref="GI:21905477"

/product="antiphagocytic M protein, type 3"

/protein\_id="AA080334.1"

/db\_xref="GI:21905477"

/translation="MAXNNTRHYSRLKLTGTASVAVALTVLGTGLVAGQTVKADAR

SVNGEPFRVKKLNEIENLLDQVTLTKHNSNYQYNAQAGRLDRQKAYLKLGLND

WAERLLQELNGEDVKKVKAPEKDDLEKVEKELDKKEKVEQDLDPKDFDLAKQ

GYVLSDRHQQLEKEKVTATKVGQISELETVKOKVETMDLTQKQNRVSLQ

EQBLATTKNAKEDFELALANADKQLEAKIADLETKLKEKEFELALGHQNAH

NEYQAKLAEKQDQIKQLEBQKQILDSARSGTARDLEAVRQAKKATSEALNNLKAELAK

VTEQKQILDASRRKGTARDLEAVRQAKQAEALQLEEQNRISSEASRRKGLRRDLASR

```

EAKQVQKDLANLTAELKVKEKQISDASRQGLRRDLDAAREAKKQVEKALBEANSK
LAALKEKLEESKLETEKAEKLEKLEAEKALAEKLEKLEKLEKLEKLEKLEKLEK
QIPDTKPGNKAIVPGKQAPQAGTTPNQKAPKMKETKPKLPSTGETANPPTFAALTMV
ATAGVAAVVKKEEN"
complement (5970..7580)
/gene="mga"
/note="synonym: SpyM3 1728"
complement (5970..7580)
/gene="mga"
/note="best blastp hit: gb|AAA26931.1| (M58461) M protein
trans-acting positive regulator [Streptococcus pyogenes]"
/codon_start=1
/transl_table=1
/product="M protein trans-acting positive regulator,
multi-gene activator"
/protein_id="AAM80335.1"
/db_xref="GI:21905478"
/translation="MYVSKLFTSQWRELKLSYVTENADAIGVKDELKALNISML
TLQCLTNQFMKEVGITYKNGIITWYHOGCLEQVQKALRHSQSFKLETLFFER
DFNSLEBELFVSLTKRLIKTNAYLHMTFGITILTSPVQVSDHQLRLEFLK
YFGEAYKISWPFCEILNKNCELLSLMKEDVVRNFTLQHLKILSSNLIYYK
GSAVDNKKTSRFSLOIQLSLQSLRFLKGLYLDLDTTAEFNSHNVDQLE
IGVAFDSIKODSPGCRKVTNWHLTNNVHLLDELEIRLSVTKRYEVAVLHNTTV
YAEITWENSFLVQNDKIRLLVIERFSVGNPKYKKEPTEPSTELTQLI
DLSEIKQDVIIVTMVGKSDLEIFFYFKMPEAIDIKLNAFLNISADSLSLDPK
IKNPLDFHRKEVITPPTPKLHAPPSTS"
7968..8162
/gene="SpyM3 1729"
7968..8162
/gene="SpyM3 1729"
/note="best Blastp hit: gb|AAF99701.1|AF276241.1
(AF276241) aminopeptidase-N [Epiphyas postvittana]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAM80336.1"
/db_xref="GI:21905479"
/translation="MTWVQPLKILSYFSQFTPTLALENVSTMFMMNLTLCLSLAVL
RQKGFVYFAFNFSHTLTS"
complement (8268..8519)
/gene="SpyM3 1730"
complement (8268..8519)
/gene="SpyM3 1730"
/note="best Blastp hit: gb|AAK34696.1| (AE006624)
hypothetical protein (mga-associated) [Streptococcus
pyogenes M1 GAS], and gb|AAL98548.1| (AE010111)
hypothetical protein [Streptococcus pyogenes MGAS8232]"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAM80337.1"
/db_xref="GI:21905480"
/translation="MIRKVTTPSQKTRKRVNGYLLKLGTAALLSLSYGILGLQP
SWENTFMGIASVAMLGVCFFIIFALNRIFDALEDNLRD"
complement (8597..10225)
/gene="isp.1"
/note="synonym: SpyM3 1731"
complement (8597..10225)
/gene="isp.1"
/note="best blastp hit: gb|AAK34697.1| (AE006624)
immunogenic secreted protein precursor [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=1
/product="immunogenic secreted protein precursor"
/protein_id="AAM80338.1"
/db_xref="GI:21905481"
/translation="MKRKLAVLLTLLISNAVPLVAVDTSLRNSTSSTDPQTAD
TDTDESETAKDKKSKETASQHDQKHPSNNHTPPSNDTKODQASSEATKPN
KKNNDTKQPNSSQSTPSKQSSQSKQSKQKQRPSPDQKQDTPKTPKRSADTK
PEKGPKEATKTEPEPNRDAKPIQPLAAAPVAPWRSDKSLKLPSSRSSAAVY
RHWTGDSAYTHNLLSRRYGTAEQLDGLFNLGSIHYDKERLNGKRLLEKLTGLDVR

```

```

AIVATAMAESSLTQGVAKKGANMFGVGFDPNPNNAKKYSDEVAIRHVVEDTIIAN
KNQIFERDLAKKWSLGOLDTLIDGVYFTDTSQSGRRADIMTKLDQWIDHGGSTP
EIPHLKITSTQSEVPVGYKRSQPNVLITYSETYSFOCTWYANRYKELGYQVD
RYMNGGDMQRKPGFVTHKPKGVYVVFAPQAGADATYGHVAVVEQIKEDSGSILIS
ESNVMLGTTISYRTTAEQASLLTYVVGDKLPRP"
complement (10327..11715)
/gene="SpyM3 1732"
complement (10327..11715)
/gene="SpyM3 1732"
/note="best blastp hit: gb|AAK34698.1| (AE006624) putative
histidine kinase [Streptococcus pyogenes M1 GAS]; similar
to two-component sensor histidine kinase"
/codon_start=1
/transl_table=1
/product="putative histidine kinase"
/protein_id="AAM80339.1"
/db_xref="GI:21905482"
/translation="MKLIKKTFIVINGLIIVVVTSLVLVYFAMPYIYTVKVKDEKVC
EFQTSQIKGKTVTIRDILTKKINDNIWYSLVSDNQLLYPSQLLDGVSEKDS
QNVNIVTTFDSYSNVKNVMSQVTLRDGKKMTLLGSSLPQVPTDASKVLLDLYPSLI
FVTVGSIIVAYLYSRTSSRILMSQTAKKMWNLENLTCTIHGKDEIAMLASDINRL
YASLTSIKSLQKEYEKASDREKSEFLRMTSHELKTPITSVIGMIDGMLNVGDF
DRKYLKRCRDVLEGAQVLSLSKLETLASQNELFSLSKLEEMEVFLVSE
LKHLYTNLEBQFVANKYLLKLNIIIDNAPHYTKSGQVMTQKDNQIVIKNEA
ETLLTQQMKQLFQPPFYRDPYSRNRKDGCTGLGLFTIHOILDQHHLAYRFVVDQRW
VFTIDFPSSHDD"
complement (11712..12365)
/gene="SpyM3 1733"
complement (11712..12365)
/gene="SpyM3 1733"
/note="best Blastp hit: gb|AAK34699.1| (AE006624) putative
two-component response regulator [Streptococcus pyogenes
M1 GAS], and gb|AAL98551.1| (AE010111) putative
two-component response regulator [Streptococcus pyogenes
MGAS8232]"
/codon_start=1
/transl_table=1
/product="putative two-component response regulator"
/protein_id="AAM80340.1"
Query Match 96.7%; Score 1046.8; DB 1; Length 53291;
Best Local Similarity 98.8%; Pred. No. 1.3e-224;
Matches 1065; Conservative 0; Mismatches 12; Indels 1; Gaps 1;
QY 1 GACAAAGCGCTCTCTTTTCTCTTACTATCTCTTAAATTTTCATATTTTAAATAAAC 60
Db 23399 GACAAAGCGCTCTCTTTTCTCTTACTATCTCTTAAATTTTCATATTTTAAATAAAC 23341
QY 61 TATTGATAAAGCTAGTAAAGTAAAGCTATCTATCTAGTTAGTAAAGTAAAGTAAAGG 120
Db 23340 TATTGATAAAGCTAGTAAAGTAAAGCTATCTATCTAGTTAGTAAAGTAAAGG 23281
QY 121 ACAACATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAATAATGTCGGCTAGTAA 180
Db 23280 ACAACATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAATAATGTCGGCTAGTAA 23221
QY 181 AATTTTCAATGGTAGTCTTCTGTATCAGCCCAAGTGGCTGTAAACACAGTCACTTGA 240
Db 23220 AATTTTCAATGGTAGTCTTCTGTATCAGCCCAAGTGGCTGTAAACACAGTCACTTGA 23161
QY 241 ATACTGACTGGCAGCAACACAGTCTCAATATGATGTTGTTCTTAAATGATGGCGCAA 300
Db 23160 ATACTGACTGGCAGCAACACAGTCTCAATATGATGTTGTTCTTAAATGATGGCGCAA 23101
QY 301 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCTAATATTACAAA 360
Db 23100 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCTAATATTACAAA 23041
QY 361 CTTTAGGTACTAGTACAGATTACTCCAGCACTCTTCTCTTAAAGCAGGAGATTTCTCTATA 420
Db 23040 CTTTAGGTACTAGTACAGATTACTCCAGCACTCTTCTCTTAAAGCAGGAGATTTCTCTATA 22981
QY 421 GCAAAATTAGTATGATAGGAGGACCGGCTACTCTAGAGGTACATTTGACTTTATGCCAATG 480

```

Db 22980 GCAAATTAGATGAGTTAGGAAGGCGGTACTGCTAGAGTACATGACTTATCCCAATG 22921  
Qy 481 TTGAAGGTAGTACGCGTGTAGACAATCTTTCGGTAAAAATCAAAACCCCGCAGATGGA 540  
Db 22920 TTGAAGGTAGTACGCGTGTAGACAATCTTTCGGTAAAAATCAAAACCCCGCAGATGGA 22861  
Qy 541 CTGGAACCCCTAATCATCTGTCAAATATAAAATGAATGTTAAATGGTCTATCTTATGTCG 600  
Db 22860 CTGGAACCCCTAATCATCTGTCAAATATAAAATGAATGTTAAATGGTCTATCTTATGTCG 22801  
Qy 601 GAGATTTCTGGAATAGATGATCTCTATGTCAGATGCTCGGTGGAGATGCACCTCAGAG 660  
Db 22800 GAGATTTCTGGAATAGATGATCTCTATGTCAGATGCTCGGTGGAGATGCACCTCAGAG 22741  
Qy 661 TCAATGCGGTTACAGGAACAGTACCCCAAAATGTAGAGGTCGTGACCAAAAAGCGCGCA 720  
Db 22740 TCATGCGGTTACAGGAACAGTACCCCAAAATGTAGAGGTCGTGACCAAAAAGCGCGCA 22681  
Qy 721 TGCCTTATACCGAACAAGAGCTCAAGATGTTAGAGCAAAATCGTGTATGCTTATCTTT 780  
Db 22680 TGCCTTATACCGAACAAGAGCTCAAGATGTTAGAGCAAAATCGTGTATGCTTATCTTT 22621  
Qy 781 ATTATGAGTCGCTCCCAATCTACACGACGAGTTCATCCAGAGCTGTCGTGCTAT 840  
Db 22620 ATTATGAGTCGCTCCCAATCTACACGACGAGTTCATCCAGAGCTGTCGTGCTAT 22561  
Qy 841 CAATGCAATCTTCGTAAATACCAATCAACGAGAAGTATTAGTTTACACACACACTAATG 900  
Db 22560 CAATGCAATCTTCGTAAATACCAATCAACGAGAAGTATTAGTTTACACACACTAATG 22501  
Qy 901 GCTACACATTAATCACTACCAATCAACGAGTACCTCAAAATATATACCAAAAGCTAGAC 960  
Db 22500 GCTACACATTAATCACTACCAATCAACGAGTACCTCAAAATATATACCAAAAGCTAGAC 22441  
Qy 961 CTCTGCTACATGAGCTAGCTTTTATCATCAAAAAGCAATGACTATAGAACTAAAA 1020  
Db 22440 CTCTGCTACATGAGCTAGCTTTTATCATCAAAAAGCAATGACTATAGAACTAAAA 22381  
Qy 1021 TACTAGAAAAGCAATGATTCGCTCATGCTTTTATGATTTTGTCAAAAAGCAAA 1078  
Db 22380 TACTAGAAAAGCAATGATTCGCTCATGCTTTTATGATTTTGTCAAAAAGCAAA 22323

## RESULT 4

AP005146/C 323825 bp DNA linear BCT 13-JUN-2003  
LOCUS Streptococcus pyogenes SSI-1 DNA, complete genome, section 6/6.  
DEFINITION AP005146 BA000034  
ACCESSION AP005146.1 GI:28811737  
VERSION  
KEYWORDS

SOURCE Streptococcus pyogenes SSI-1  
ORGANISM Streptococcus pyogenes SSI-1  
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
Streptococcus.

REFERENCE 1  
AUTHORS Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y.,  
Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T.,  
Hayashi, H., Hattori, M. and Hamada, S.  
TITLE Genome Sequence of an M3 Strain of Streptococcus pyogenes Reveals a  
Large-Scale Genomic Rearrangement in Invasive Strains and New  
Insights into Phage Evolution  
JOURNAL Inmate Res. 13 (6), 1042-1055 (2003)  
MEDLINE 22683278  
REFERENCE 2 (bases 1 to 323825)  
AUTHORS Yamashita, A., Nakagawa, I., Kurokawa, K., Nakata, M., Tomiyasu, Y.,  
Yamazaki, K., Okahashi, N., Kawabata, S., Yasunaga, T., Hattori, M.,  
Hayashi, H. and Hamada, S.  
TITLE Direct Submission  
JOURNAL Submitted (01-MAY-2002) Ken Kurokawa, Osaka University, Genome  
Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,  
Japan (E-mail: ken@gen-info.osaka-u.ac.jp,  
URL: http://www.gen-info.osaka-u.ac.jp/, Tel: 81-6-6879-8365,

COMMENT genome project  
This clone was isolated from a patient presenting with toxic shock  
like syndrome.

FEATURES  
source

1..323825  
/organism="Streptococcus pyogenes SSI-1"  
/mol\_type="genomic DNA"  
/strain="SSI-1"  
/db\_xref="taxon:193567"  
complement (422..1012)  
/gene="Spe1572"  
complement (422..1012)  
/gene="Spe1572"  
/notes="similar to GB:AAK33432.1 (AE006502) percent  
identity 100 in 196 aa"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative ATP-dependent protease proteolytic  
subunit"  
/protein\_id="BAC64667.1"  
/db\_xref="GI:28811738"  
/translation="MIPVIEOTSRGERSVDIYRLLKDRILMLTGPVEDNMANSVIA  
QLFLDAQNTKDIYLYVNTPGSVSAGLAIVDTMFIKADVQTIWGMMAASMGTVIA  
STQTKGRFPLNAEYMIHOPMGGTGGTQQTDMATAAEHLKTRHRLKILQNAQK  
TIQIKHDAERDYWMAEETLAYGFIDEIMNNELK"  
complement (1235..1864)  
/gene="Spe1573"  
complement (1235..1864)  
/gene="Spe1573"  
/notes="similar to GB:AAK33430.1 (AE006502) percent  
identity 100 in 209 aa"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="putative uracil phosphoribosyltransferase"  
/protein\_id="BAC64668.1"  
/db\_xref="GI:28811739"  
/translation="MGKCOVISHPLIQHKLILRRQTTSTKDFRELVNEIAMLMGVY  
SRDPLEVDIOTPVSKTQKOLAGKLAIVPILRAGIGWVGGLSLVPAKVGHIQW  
YRNEETLEPEYLVKLPELINQRIQFVDPMLATGSAIILAVSLKRGANIKFVCL  
VAAPEGVKQLQEAHPDIDIFTAALDDHLEHGYIVFGLGAGDRFLPGTK"  
complement (2032..3666)  
/gene="Spe1574"  
complement (2032..3666)  
/gene="Spe1574"  
/notes="similar to GB:AAK33429.1 (AE006501) percent  
identity 99 in 544 aa"  
/codon\_start=1  
/evidence=not\_experimental  
/transl\_table=11  
/product="conserved hypothetical protein"  
/protein\_id="BAC64669.1"  
/db\_xref="GI:28811740"  
/translation="MSTERRKQLQEEELMVQGAWSAGTAFISRLGLVLYIPWYMWG  
QVAIQALFNMGYNVYVFLISTTGLNVAIAKQVAKNMGQTHSYQLIRSTLKL  
MLGLHIFSAIYVLGSLPASISGDDTLVPIMHSLSLAVFIIPVMSVIRGIPQHNH  
IKPYAISQIAEQLIRVIMLLTFLIMKLGSDYASAVTQSTFAAIFGMVSMGLVY  
YLMKQGLAARAKSPDHTSIDINGLLETLESIFIVTGSALIQFOLLIDQWTFVNT  
MTLTDYSRQLLVLFYFNANPAKITVLIIVAAISIGGVGIALTFENYKKDKMAA  
RLTIINNIEMLVFALPTGATILARPLSVFYFYGASEDRAIHLFVAVLFTLLMALY  
LFSPLQALFENKAIYFAYGILIKLQIPLIYLLHAYGLPIATIALVPIYLMY  
RLYQVTHFNRLKQLRLTLLETLLMGLVFWANLLGYAFKPTGRITSLLVLLII  
GGLGMVYVYATLTLLTQDLKLGSKASRLRQLKLGWH"  
3754..5199  
/gene="Spe1575"  
3754..5199  
/gene="Spe1575"  
/notes="similar to GB:AAK33428.1 (AE006501) percent  
identity 99 in 481 aa"  
/codon\_start=1  
/evidence=not\_experimental

```

/transl_table=11
/product="putative UDP-N-acetyl-muramoylalanine-D-glutamyl-2-
6-diaminopimelate ligase"
/protein_id="BAC64670.1"
/db_xref="GI:28811741"
/translation="MTIEQLDILIKKHNFREVLDADEYHYHQVSPERLSYDSRQ
VFGKTLFPKATFKADYKLEAIKNGQLYISEVDYELGIPVVLVTDIKKAMSLIAMA
TPESLDLPFAMAEVCVTGMWHLIMEVSSQAYLVDRVYGLTFDVGVLNTPSPDHGPIE
HPTPEDEVYHKKLLMENSRAVINSVMDHESFLADQVADQEHFYFGLSDNQITTSQA
ESPAKQGLAHDYDIQLIGHFNQENMAAGLACILGASLADIQKIGAKTRVPRGMEV
LIMTHAKVFDYAHNGDSLEKLSVVEEHTQKMLILGAPGNKGSRRADPRGVH
QHPNLWILTADNFDEPDIDISKEIASHARVPEIISDREQAIQKMSLCQAKDAV
IAGKGADAYQIVKGGQVAVAGDLAIKHYL"
/length=6227
/gene="SPa1576"
CDS
5445..6227
/gene="SPa1576"
/notes="similar to GB:AAK33427.1 (AE006501) percent
identity 99 in 260 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative ferrichrome ABC transporter
(ATP-binding protein)"
/protein_id="BAC64671.1"
/db_xref="GI:28811742"
/translation="MTTISAEDLTIAVEQRTIIDKLSFVIPSGKITTIIGANGCKSS
LLKALTRLLPQKGVVNGONATLETKEVAKKALLPQVQENNGITVVELSVYGR
PHQSYFNGNLPADKKAHWMQATNMAYADQVDAISGGQQRVWMLAALAOQTDI
IFDEPTVTLNHLQLEILELKSNDKAGKTIIMVLDLNLISARYSDHLIAKHKGI
HYGTIADWMSPIIDIFQIKFVLVDDPIHNCPIVLTYLQ"
/length=6253
/gene="SPa1577"
CDS
6253..7185
/gene="SPa1577"
/notes="similar to GB:AAAL97176.1 (AE009984) percent
identity 99 in 310 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative ferrichrome ABC transporter
(ATP-binding protein)"
/protein_id="BAC64672.1"
/db_xref="GI:28811743"
/translation="MKKTLLELTLCLTITLIACNQATNHSNTASKSLSPMQIAGV
TYGIDPKPRKRVSLASTYTGKLDMLNVLGVTSDKNPILAKTVKAKOVAATD
LEAITLPDLIVGSTEENIKQAEIAPVISIEYRKDYLOVLSDFGRIFNKGKAK
KWLKDWKTAAAYEKEVAVTGDKATFTIMGLYEKDVYLGKDWGRGELIHOAFHYD
APEKVTVEFKGYSLSQEVLPDYIGDYVVIADDDKTSALYESKMWQSIAPVKKH
HVIKVNANVFYTDPLSLEYQLETLREALISEN"
/length=7196
/gene="SPa1578"
CDS
7196..8227
/gene="SPa1578"
/notes="similar to GB:AAAL97175.1 (AE009984) percent
identity 99 in 343 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative ferrichrome ABC transporter (permease)"
/protein_id="BAC64673.1"
/db_xref="GI:28811744"
/translation="MCLERFCLKKTQIITKSHIFLWVLLSFLSVAIYCHIRFGAV
ALSHQDNLNLSILFGKNGHKANVLLAIELPLRGATLGSALAVSGTMAITRNPAE
PGLGINAGLALVAYAFPHLYSHIILSLGSLAATLVFGLSYQSGKGHQL
RLVLGAMVSTILLSALCGGITHLANAVIQWAGGLGVNMWQMTGYAPITILSLC
LAQLSHLTIVLSLUSEQAKQKTNLISAVFMIIVLILSSAAVAIAGSIGFGLVI
PHLMKHFTPHRYLLPLCAVSGASFMWVDIACRNLPPEYETPLGALVSLIGFPCL
WLIRRGGRY"
/length=8188
/gene="SPa1579"
CDS
8188..9225
/gene="SPa1579"

```

```

/gene="SPa1579"
/notes="similar to GB:AAK33424.1 (AE006501) percent
identity 98 in 345 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative ferrichrome ABC transporter (permease)"
/protein_id="BAC64674.1"
/db_xref="GI:28811745"
/translation="MLFMVNRKRLPKSHRHLGLYTSILILVLSLMLGLSLGESH
LSPDILHVLGKSSHAISFIVINRILPRILAACLGGSALSLGLIQLRTRNPLADS
VGLDITGAGISLAIIVVSFFEQAHISHYLPLFAMLGAVVTFVSVMLSLITQGOID
PTRILITGVAVTMLSSLMVALVGHINRYKVDLVINWLSQLIGDOWPTLSVAPALL
CFWLLTYSQAHLNIMGLADNATIGLGLPLNRKRLLIULVLAAGLSGLVNLGNSFI
GLIAGHFTSVLVGNSHKLTPISILGMILLVADTVGRVILVGSNIQTGILVSLIGA
PYFLYLMAKTK"
/complement(9270..9923)
/gene="SPa1580"
CDS
9923..10000
/gene="SPa1580"
/notes="similar to GB:AAAL97173.1 (AE009984) percent
identity 98 in 217 aa"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="BAC64675.1"
/length=96
/Query Match
Best Local Similarity 96.7%; Score 1046.8; DB 1; Length 323825;
Matches 1065; Conservative 0; Mismatches 22; Indels 1; Gaps 1;
QY 1 GACACGCTCTCTTTTCTCTTACTATCTCTCTTAAATTTTCAVATTTTAAAAAAC 60
DB 188896 GACACGCTCTCTTTTCTCTTACTATCTCTCTTAAATTTTCAVATTTTAAAAAAC 60
QY 61 TATTGATAACAGTAGTTAAGTAAGCGTATCTATGTTAGTTCGCGAAATAGGAGG 120
DB 188837 TATTGATAACAGTAGTTAAGTAAGCGTATCTATGTTAGTTCGCGAAATAGGAGG 120
QY 121 ACAAGCATATGAATCTTCTTCTGATCAAGACGGTCTTCTTAAATTTTCAVATTTTAAAAAAC 180
DB 188777 ACAAGCATATGAATCTTCTTCTGATCAAGACGGTCTTCTTAAATTTTCAVATTTTAAAAAAC 180
QY 181 AATTTTCAATGCTAGCTCTTCTTCTGATCAGGCAATCGCTGTAAACACAGTTCACCTTGAAA 240
DB 188717 AATTTTCAATGCTAGCTCTTCTTCTGATCAGGCAATCGCTGTAAACACAGTTCACCTTGAAA 240
QY 241 ATACTGCACTGCGACGACAAACACACAGGCTCAAAATGATGTTGTTCTAAATGATGCGCGAA 300
DB 188657 ATACTGCACTGCGACGACAAACACAGGCTCAAAATGATGTTGTTCTAAATGATGCGCGAA 300
QY 301 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGCTCACTTAACTATTACAAA 360
DB 188597 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGCTCACTTAACTATTACAAA 360
QY 361 CTTTAGGTACTAGTCAGATTACTCCAGCCTCTTCTTCTTAAAGCAGGAGATATCTCTATA 420
DB 188537 CTTTAGGTACTAGTCAGATTACTCCAGCCTCTTCTTCTTAAAGCAGGAGATATCTCTATA 420
QY 421 GCAATTAGAGTAGTTAGGAGGCGGTACTGCTAGAGGTACATTCAGCTTATGCCAATG 480
DB 188477 GCAATTAGAGTAGTTAGGAGGCGGTACTGCTAGAGGTACATTCAGCTTATGCCAATG 480
QY 481 TTGAAGGTAGCTAGCGTGTAGCAATCTTTCGTTAAATTCAAAACCCCGCAGGATGGA 540
DB 188417 TTGAAGGTAGCTAGCGTGTAGCAATCTTTCGTTAAATTCAAAACCCCGCAGGATGGA 540
QY 541 CTGAAAACCCCTAAATCATGTCAAAATATAAAATGAAATGTTAAATGTTCTATCTTATGTCG 600
DB 188357 CTGAAAACCCCTAAATCATGTCAAAATATAAAATGAAATGTTAAATGTTCTATCTTATGTCG 600
QY 601 GAGATTTCTGGAATAGAGTCATCTCATTCGAGATAGTCTCGGTGGAGATGCACTCAGAG 660

```

```
Db 188297 GAGATTTCTGGAATAGAGTCAATCTCATTTGAGATAGTCTCGGTGGAGATGCATCTCAGAG 188238
Qy 661 TCAATGCCGTTTACAGGAACACGTACCCAAAATAGTAGGAGTCTGTCACCAAAAAGCGCGCA 720
Db 188237 TCAATGCCGTTTACAGGAACACGTACCCAAAATAGTAGGAGTCTGTCACCAAAAAGCGCGCA 188178
Qy 721 TGCCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780
Db 188177 TGCCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 188118
Qy 781 ATTATGAAGTCGCTCCCAATCTACCAACGACGAGTGTGATTCCAAGAGCTGCTGCTGAT 840
Db 188117 ATTATGAAGTCGCTCCCAATCTACCAACGACGAGTGTGATTCCAAGAGCTGCTGCTGAT 188058
Qy 841 CAATGCAATCTTCTGATAATACCAATCAACGAGAAAGTATTAGTTTACAACACACTAATG 900
Db 188057 CAATGCAATCTTCTGATAATACCAATCAACGAGAAAGTATTAGTTTACAACACACTAATG 187998
Qy 901 GCTACACCAATTAACCTACCAATACCGGTACACCTACTCAAAAATATATACCAAAAGCTAGAC 960
Db 187997 GCTACACCAATTAACCTACCAATACCGGTACACCTACTCAAAAATATATACCAAAAGCTAGAC 187938
Qy 961 CTCTGCTACCTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAAA 1020
Db 187937 CTCTGCTACCTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAAA 187878
Qy 1021 TACTAGAAAAAGCAATGATTCGCTGCTGCTTTTATGATTTTATGATTTGTCAAAAAGCAAA 1078
Db 187877 TACTAGAAAAAGCAATGATTCGCTGCTGCTTTTATGATTTTATGATTTGTCAAAAAGTAAAA 187820

RESULT 5/
LOCUS AE006625/c 12808 bp DNA linear BCT 01-JUN-2001
DEFINITION Streptococcus pyogenes M1 GAS strain SF370, section 154 of 167 of
the complete genome.
ACCESSION AE006625 AB004092
VERSION AE006625.1 GI:13623035
KEYWORDS
SOURCE
ORGANISM Streptococcus pyogenes M1 GAS
Streptococcus pyogenes M1 GAS
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 12808)
AUTHORS Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K.,
Primeaux,C., Szate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S.,
Qian,X., Jia,H.G., Najjar,F.Z., Ren,Q., Zhu,H., Song,L., White,J.,
Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.
Complete genome sequence of an M1 strain of Streptococcus pyogenes
Proc. Natl. Acad. Sci. U.S.A. 98 (8), 4658-4663 (2001)
JOURNAL 21192684
MEDLINE
PUBMED 11296296
REFERENCE 2 (bases 1 to 12808)
AUTHORS Ferretti,J.J., McShan,W.M., Adjic,D., Savic,D., Savic,G., Lyon,K.,
Primeaux,C., Szate,S.S., Surorov,A.N., Kenton,S., Lai,H., Lin,S.,
Qian,X., Jia,H.G., Najjar,F.Z., Ren,Q., Zhu,H., Song,L., White,J.,
Yuan,X., Clifton,S.W., Roe,B.A. and McLaughlin,R.E.
Direct Submission
Submitted (10-APR-2001) Department of Microbiology and Immunology,
University of Oklahoma Health Sciences Center, 940 SL Young Blvd,
Oklahoma City, OK 73104, USA
JOURNAL
PUBMED
FEATURES
Source
1. .12808
/organism="Streptococcus pyogenes M1 GAS"
/mol_type="genomic DNA"
/strain="SF370"
/serotype="M1"
/db_xref="taxon:160490"
complement(242..646)
/gene="SPY2033"
complement(242..646)
/gene="SPY2033"

gene
CDS
```

```
/note="Best Blastp hit = pir||T27702 hypothetical protein
zk1151.3 - Caenorhabditis elegans
>gi|3881532|emb|CAB07723.1| (293398) cDNA EST yk4e8.3
comes from this gene
cDNA EST yk4e8.5 comes from this gene
cDNA EST EMBL:T00673 comes from this gene
cDNA EST yk20d2.5 comes from this gene
cDNA EST yk2398.5 comes from this gene
cDNA EST yk37e2.5 comes from this gene
cDNA EST>
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK34703.1"
/db_xref="GI:13623036"
/translations="MTGQKKLANLKKTKQIIIVRTVDASVKLKVSQKQEHFIMKFK
KLVIVPALALAAATCFLTACGTGKDSKEVKEIKMSDIKDDAVSKTKTVVDGEVTEY
TTKDGNVIOIPAGNEEGMESKADGGGAPAKN"
complement(673..969)
/gene="SPY2034"
complement(673..969)
/gene="SPY2034"
/note="Best Blastp hit = no blast hits"
/codon_start=1
/transl_table=11
/product="conversed hypothetical protein"
/protein_id="AAK34704.1"
/db_xref="GI:13623037"
/translations="MTKQDOLIVKMEQTEAFSPKLANLIEALDAFKEHVEEYATLR
NFYSSDEWFLANQPWDDIPCGVLSDDLFDIMIGHNQLLADILDALPIYKHM"
complement(1725..2654)
/gene="SPY2037"
complement(1725..2654)
/gene="SPY2037"
/note="Best Blastp hit = sp|P15294|PRTM_LACLA PROTEASE
MATURATION PROTEIN PRECURSOR >gi|79978|pir||S08083
probable protein export protein prtm precursor - 763)
Lactococcus lactis subsp. cremoris (strain NCDO 763)
plasmid pLP763 >gi|47198|emb|CAA32349.1| (X14130) ORF (AA
1 to 299) [Lactococcus lactis subsp. cremoris]"
/codon_start=1
/transl_table=11
/product="conserved hypothetical"
/protein_id="AAK34705.1"
/db_xref="GI:13623038"
/translations="MKQMNKLIITVTVTLATVTLTACQSHNNTKLYSMKGDITITVD
FYNKNTLTAQKAMLSLVTSRVFETQYANKVSKDEKAYKOTADQYGTFSFKTVLAQ
SGLPTPKYQIRLTKLVEYAVKEAQNETISKDYQAYDAYTPTWTAIEMQPEKEE
DAKALEAVKAGADFAAIAEKRTAADRKTITTFDSGETILPAEVVRAASGLKEGKR
SEITLALDPATSKRTYHIITKVTTKATKADKAYQKRLKDIIVTGKLDKDPDFQNKVIA
KALDKANVKIKDAFANILAQFAKPNQKQPAQR"
complement(3134..3439)
/gene="SPY2038"
complement(3134..3439)
/gene="SPY2038"
/note="the gene fragment has weak similarity to amino
acids 22-128 of pyrogenic exotoxin B; Best Blastp hit =
gb|AAA26986.1| (L26157) pyrogenic exotoxin B
[Streptococcus pyogenes]"
complement(3441..4637)
/gene="speB"
/note="synonym: SPY2039"
complement(3441..4637)
/gene="speB"
/note="Best Blastp hit = gb|AAA26982.1| (L26153) pyrogenic
exotoxin B [Streptococcus pyogenes]"
>gi|431624|gb|AAA26984.1| (L26155) pyrogenic exotoxin B
[Streptococcus pyogenes] >gi|431650|gb|AAA26997.1|
(L26131) pyrogenic exotoxin B [Streptococcus pyogenes]
>gi|431664|gb|AAA27004.1| (L26138) pyrogenic exotoxin B
[Streptococcus pyogenes] >gi|431688|gb|AAA27016.1|
(L26150) pyrogenic exotoxin B [Streptococcus pyogenes]"
/codon_start=1
```

```
/transl_table=11
/product="pyrogenic exotoxin B"
/protein_id="AAK34706.1"
/db_xref="GI:13623039"
/translation="MNKKLGLVRLSLALLGPFVLANPVFADQNFARNEKEAKDSAIT
FIQGAALAGARSADIKLQKVALGELSGSNMYVNIITGGFVIVSGDKRSPKILG
YSTGSPFANGKNIASFMESYVQIKENKLDITTYAGTAIEIKOPVVKSLDKGIHY
NQGNFYNLITFVIEKVPGEQSFVGOHAATGCATATAQAIKMYHNPYKGLDKYITTL
SSNPFYHNPKNLFAAISTROYNNLPTYSRESNVQKMAISELMVGISVDMYD
GPSGAGSSRVORALKENFCYNSVHQINRGFSKQDWEAQIDKELSONOPVYQGV
KVGSHAFVIGAGGRNFYHVNWGVGSDGFFRLDALNPSALGTGGGAGNGYQSA
VVGKIP"
complement (4909..5079)
/gene="SPY2040"
CDS
complement (4909..5079)
/gene="SPY2040"
/notes="Best Blastp hit = no blast hits"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK34707.1"
/db_xref="GI:13623040"
/translation="MDNPKDKGLSLGVDPITILITESVINKMRTVPVHLKLSST
YCLLDLAYOSA"
5163..5234
/gene="SPY2041"
CDS
5163..5234
/notes="Best Blastp hit = no blast hits"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK34708.1"
/db_xref="GI:13623041"
/translation="MLISVDKLNKLNKMLLFVNAIDFSLSTILQKELQK"
5578..6420
/gene="ropeB"
CDS
5578..6420
/notes="synonym: SPY2042"
/gene="ropeB"
/notes="Best Blastp hit = dbj|BAB16026.1| (AB030578) ropeB
[Streptococcus pyogenes]"
/codon_start=1
/transl_table=11
/product="putative transcription regulator"
/protein_id="AAK34709.1"
/db_xref="GI:13623042"
/translation="MEIGETVEFIRHSKNSIKQVGDYLTROTYYRPIKNNLDISK
KLXYILDNLNVNDFLISNNFQYKEFIDMTAKHYFEGRNIEGLNHILDSYKDSK
STKENLFLAVKVLATLTERDCLTERTYLSNVLINLETWSHETVLFNKMCFPESC
FIEMVFSKVLNLDKYNTRYGNESIRMFVNMILLFQROYDKASEILLAKIEDYOL
NDCLYERCCVSFFDGIIGLINGEGAEQKVCQILEIFQLNCKTIHMFQTYLEAIK
HKLS"
complement (6661..7476)
/gene="mf"
CDS
/notes="synonym: SPY2043"
complement (6661..7476)
/gene="mf"
/notes="Best Blastp hit = emb|CAA59264.1| (X84793)
streptodornase [Streptococcus pyogenes]"
/codon_start=1
/transl_table=11
/product="mitogenic factor"
/protein_id="AAK34710.1"
/db_xref="GI:13623043"
/translation="MNLIGSRVFSKCRILKFSMVALVSATMAVTVTLTENTLARQ
TQVNSDVLNDGASKYLNEALWTFDPSNYKLTGTSQITPALFPKAGDILYSKIDE
LGRNTRARGITLTVSGSYGVQSGFNQNPAGTGNPNHVKYKLEWNLGSLSYGDF
WNRHLLADSLGGDALRVNATGTRTQNGVRDQKGMRYTEORAOEWLEARNRGLY
YEAPFIYNADLPIPRVVVSMQSSDNTINEKLVYNTANGTYTINHNGPTQK"
7840..8349
/gene="SPY2045"
CDS
7840..8349
```

```
/gene="SPY2045"
/notes="Best Blastp hit = sp|O32124|YUTG_BACSU_HYPOTHETICAL
18.7_KDA_PROTEIN_IN_HOM-MRGA_INTERGENIC_REGION
>gi|7474749|pir||A70024 conserved hypothetical protein
yutG - Bacillus subtilis >gi|2635725|emb|CAB1518.1|
[299120] similar to hypothetical proteins from B. subtilis
[Bacillus subtilis]"
/codon_start=1
/transl_table=11
/product="protein low temperature requirement C"
/protein_id="AAK34711.1"
/db_xref="GI:13623044"
/translation="MSTDHQLREVSYOLLAERGVSLEIAELVLFLONDYIPNLTMRE
CLESEAVLAKREVQNAIITGVELDKLAENQISEPLLSILKTDQGLYDEILALSI
VNLVSGTGFTNYGLDKTKPSIVDKLNKDGYSCHTFLLDDIVSAIAAASRIAHNDP
96.4%; Score 1044.4; DB 1; Length 12808;
Query Match
Best Local Similarity 98.1%; Pred. No. 4.6e-224;
Matches 1057; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 1 GACAAACGCTTCTTTTCTCTTACTATCTCTCTTAAATTTTTCATATTTTAAAAAAC 60
DB 7604 GACAAACGCTTCTTTTCTCTTACTATCTCTTAAATTTTTCATATTTTAAAAAAC 7545
QY 61 TATTGATAAACTAGTTAAGTAAGCGGTATATCTATGTTAGTTCGAAATTAGAAAAAGG 120
DB 7544 TATTGATAAACTAGTTAAGTAAGCGGTATATCTATGTTAGTTCGAAATTAGAAAAAGG 7485
QY 121 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGTAGTAA 180
DB 7484 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGTAGTAA 7425
QY 181 AATTTCATPGTAGTCTTGTATACGACCAATGGCTGTAAACAGTCACACTTTGAA 240
DB 7424 AATTTCATPGTAGTCTTGTATACGACCAATGGCTGTAAACAGTCACACTTTGAA 7365
QY 241 ATACTGCACCTGGCACCACAAACACAGGTCTCAATGATGTTGTTCTAAATGATGCGCAA 300
DB 7364 ATACTGCACCTGGCACCACAAACACAGGTCTCAATGATGTTGTTCTAAATGATGCGCAA 7305
QY 301 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCACTAATATTACAAA 360
DB 7304 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCACTAATATTACAAA 7245
QY 361 CTTTAGTACTAGTCAGATTACTCCAGACCTCTTTCCTAAAGCAGAGATATCTCTATA 420
DB 7244 CTTTAGTACTAGTCAGATTACTCCAGACCTCTTTCCTAAAGCAGAGATATCTCTATA 7185
QY 421 GCAATTAGATGAGTTAGGAGGACCGCTACTCTGCTAGAGGTACATTGACTTATGCCAATG 480
DB 7184 GCAATTAGATGAGTTAGGAGGACCGCTACTCTGCTAGAGGTACATTGACTTATGCCAATG 7125
QY 481 TTGAAGTAGTACCGGTGTTTACAAATCTTTTCGTTAAATAATCAAAACCCCGCAGATGGA 540
DB 7124 TTGAAGTAGTACCGGTGTTTACAAATCTTTTCGTTAAATAATCAAAACCCCGCAGATGGA 7055
QY 541 CTGGAACCCCTAAATCATGTCATAATATAAATTAATGATGTTAAATGTTCTATCTTATGTCG 600
DB 7054 CTGGAACCCCTAAATCATGTCATAATATAAATTAATGATGTTAAATGTTCTATCTTATGTCG 7005
QY 601 GAGATTTCTGGAATAGAGTCAATCTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAG 660
DB 7004 GAGATTTCTGGAATAGAGTCAATCTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAG 6945
QY 661 TCAATGCGGTTACAGGACCGTACCCCAAAATGATGAGGTGCTGACCAAAAAGCGGCA 720
DB 6944 TCAATGCGGTTACAGGACCGTACCCCAAAATGATGAGGTGCTGACCAAAAAGCGGCA 6885
QY 721 TGGCTATACCGAAACAAAGAGCTCAAGATGTTTAAAGCAAAATCGTATGCGGTCTCTTT 780
DB 6884 TGGCTATACCGAAACAAAGAGCTCAAGATGTTTAAAGCAAAATCGTATGCGGTCTCTTT 6825
QY 781 ATTATGAGTCGCTCCAAATCTACACGACAGAGTTGATTCACAGAGTGTCTGCTGAT 840
```





```

QY 961 CTCTGCTACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAA 1020
Db 1646 CTCTGCTACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAA 1705
QY 1021 TACTAGAAAAACAATGATTGCCGTCATTGCTTTTATGATTTTGTGCAAAAGCAAA 1078
Db 1706 TACTAGAAAAACAATGATTGCCGTCATTGCTTTTATGATTTTGTGCAAAAGTAAAAA 1763

RESULT 7
LOCUS AE010113/c
DEFINITION Streptococcus pyogenes strain MGAS8232, section 161 of 173 of the
complete genome.
ACCESSION AE010113 AE009949
VERSION AE010113.1 GI:19749176
SOURCE Streptococcus pyogenes MGAS8232
ORGANISM Streptococcus pyogenes MGAS8232
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
1 (bases 1 to 10424)
Smoot, J.C., Barbican, K.D., Van Gompel, J.J., Smoot, L.M.,
Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M.,
Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M.,
Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G. and Musser, J.M.
Genome sequence and comparative microarray analysis of serotype M18
Group A Streptococcus strains associated with acute rheumatic fever
outbreaks
Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4668-4673 (2002)
2 (bases 1 to 10424)
Smoot, J.C., Barbican, K.D., Van Gompel, J.J., Smoot, L.M.,
Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M.,
Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M.,
Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G. and Musser, J.M.
Direct Submission
Submitted (31-JAN-2002) Laboratory of Human Bacterial
Pathogenesis/Rocky Mountain Laboratories/NIH, 903 S. 4th St.,
Hamilton, MT 59840, USA
Location/Qualifiers
1. .10424
/organism="Streptococcus pyogenes MGAS8232"
/mol_type="genomic DNA"
/strain="MGAS8232"
/db_xref="taxon:186103"
/complement(107..922)
/gene="spyM18_2104"
/complement(107..922)
/gene="spyM18_2104"
/feature="best blastp match pir||S36907 mitogenic factor,
25K, precursor - Streptococcus pyogenes"
/codon_start=1
/transl_table=11
/product="mitogenic factor, 25K, precursor"
/protein_id="AAL98563.1"
/db_xref="GI:19749177"
/translation="MNLGSRVFSKRLVKFSWVALVSATMAVTTVLTALARQ
TQVNDVLDGAKYLINEALWTFNDSPNYKTLGTSTQITPALPKAGDILYSKLDE
LGRTRTARGTLTAYANVEGSGVRSFGKQNPAGTGNPNHVYKLEWNLGSLYVGD
WNRSHLADSGLGDALGNVAVTGRTONTGVGRQKGMRYTEORAEWLEARNRGYLY
YEAPITNADBLIPRAVWVSQSSDNTNEKVLVYNTANGTYINHGTFPTQK"
1286..1795
/gene="spyM18_2106"
1286..1795
/gene="spyM18_2106"
/feature="best blastp match gb|AAK34711.1| (Streptococcus pyogenes M1
GAS)"
/codon_start=1
/transl_table=11

```

```

/product="putative low temperature requirement C protein"
/protein_id="AAL98564.1"
/db_xref="GI:19749178"
/translation="MSTDHHLREVSQYLAERGVSLEAIAELVLFLONNYIPNLTMAE
CLESEVAALAKREVNAIITGVELDKLAERANQLSEPLLSILKTDQGIYDEIDLAISI
VNLXGSLTGLNYGLDKTKPGICVDKLNHKGKSCHTFLDDIVSAIAAAAAARIAHNDP
AKSAITNQO"
/complement(1877..2965)
/gene="glcA"
/feature="synonym: spyM18_2108"
/complement(1877..2965)
/gene="glcA"
/feature="best blastp match gb|AAK34712.1| (AE006625)"
/product="putative glycerol dehydrogenase (Streptococcus pyogenes M1
GAS)"
/codon_start=1
/transl_table=11
/product="putative glycerol dehydrogenase"
/protein_id="AAL98565.1"
/db_xref="GI:19749179"
/translation="MKVFASPSRYIQKVALFTNAKTLKQDGSPIILLCDDVVYGVIG
ERESYLIDNGMTPVHVFANGSEASNEISRVAIAKENSNDVIIGLGGKTTDSAKAI
ADLLAVPIAPTIASTADTSALSVIYDEGAFKFIYFSKNPDLVLDVTOVIQOAP
KRLASGIADGLATVWEARVMQKNGDWTAGNQTLAGVAIAKACEQTLFADGLKAMA
SCDRQVTPALENVEANTLLSGLGFESAGLAAAHAINHGFTALTGAIHLLTHGKVA
YGTQLTFLENSREEIDRYIDFYQAIGMPTTLKEMHLDATQEDFLKLGROATWAGE
TIHQMPFVISPEDVAAALVADAVYTSR"
/complement(3022..3690)
/gene="spyM18_2109"
/complement(3022..3690)
/feature="best blastp match gb|AAK34713.1| (AE006625)"
/product="putative transaldolase-like protein (Streptococcus
pyogenes M1 GAS)"
/codon_start=1
/transl_table=11
/product="putative transaldolase-like protein"
/protein_id="AAL98566.1"
/db_xref="GI:19749180"
/translation="MEYMLDTLDLEAIKKWHILPLAGVTSNPISIAKKGDDIDFFERI
REVRAIIGDKASIHVQVIAQDYGLIKLDAIRQCDSVYVKVPTTEGLAAIKTLK
ABGHYITATVYITFOGLIAEAGADVLAPYNNMENLNDPEAVIGOLAEAINRENA
DSKILAAAFKNAQVKNKSFALGAQAITAGPDVFEAGFAMPSIQKAVDDFGKOWEAIH
RKS"
/complement(3703..6120)
/gene="pfld"
/feature="synonym: spyM18_2110"
/complement(3703..6120)
/gene="pfld"
/feature="best blastp match gb|AAK34714.1| (AE006625)"
/product="putative pyruvate formate-lyase 2 (Streptococcus pyogenes
M1 GAS)"
/codon_start=1
/transl_table=11
/product="putative pyruvate formate-lyase 2"
/protein_id="AAL98567.1"
/db_xref="GI:19749181"
/translation="MTETKSPFGHLTDTRMTHREAVLDKPKYIDAERAILATEAYQK
HQKPAKKAALMTILENMTIYIEDESLIAGNOASNNKADPFPPEYTLFVLNLD
LFKRGDGVYITTEKQQLRDIAFPENNRLRQCVLLPEVQVMTPEFGMGK
MNSGDHALVNYQKLEHLGKFEERAAKALDLTIPENDKHFFYDSVFIIVDAK
KYAKYAQALARELAETASPEROELLDIACIDKVPYEPARTFAVQSVWFIOCL
QIESNGHSYGRFDQVYVYKADLEAGRETDIVERTLNLWIKTLINKVRSQAH
TFSSAGSLYQNVITIGOTRDKDANPLSYLVLSVAQTKLPQNPNTVRYHKGDN
FMNECLVMKLGFMAMPANDEIIPFIRKGVSEEDAYDYSKAAVATFVFKWGYR
CTGMSINFPKLLITMNDIDPASGRFAKHGFMDTSYEELKAAWADATLRETR
MSVINAIDLGLEREPFELWQALSDDDICGRGKTLKGGAVDYISGQVQANLSD
SLAALKVFEGRLTPEELWQALSDDFAGEREIRQMLINDAPKYQNDNDYASLV
VEAYDYIDYIRKPNYGRGPIGGIRYGTSSISANVQCKGTLATPDGRAGTFL
AECSEPHSMWKKGPTSLVSKVAKLPDDEIVGVLLNKNVNPQTLAKEDKLMALL
RTFENLGHVYIQNVVSVRETLIDAKQHEKRDILVRVAGYSFAFVLSUKATQDDII
ERTEHTL"
/complement(6331..7635)

```

gene



```

CDS
/gene="spym18_2111"
/complement(6331..7635)
/gene="spym18_2111"
/notes="best blastp match gb|AAK34715.1| (AB006626)
putative PTS system, enzyme IIC component [Streptococcus
pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative PTS system, enzyme IIC component"
/protein_id="AAL98568.1"
/db_xref="GI:19749182"
/translacion="MAKNMNCNIIMPITFVNRGIITALKDGMILAILPLTVVGLSLFLI
AGQIPFQGINADTAGVGDWTEPPMQVYHGTFAIMGLISCPALGISYAKNSGVEPLP
SGULSAPFILLRSSVPAGEFIDAIKSNKMPGGGIIIGAVIGUTVGAIYATFIR
RHVIMKPEQVQPIAKQFEAMIPAFVIFLSMLVYIIAKSVTGGSTFIEMIYDVIOV
PLOGLTGSLGALGIAFFISPLMFWGVHGSVNVGIVTALLSLNLDANKALMAAGELS
LDKGAHTVTOQFIDSLFISGSGITFGLVAMIPAAKSKKALGVKVAAPALPNVNE
PVVGFPIVMNPVMPFLIPVLAALTIVVGAIGFMQPPFAGVTLPLWSPALISGRM
VGGWGAIVQILLILINSLVYFFPKIQDNMAYQNEQASES"
/complement(7645..7953)
/gene="spym18_2113"
/notes="best blastp match gb|AAK34716.1| (AB006626)
putative PTS system, enzyme IIB [Streptococcus pyogenes M1
GAS]"
/codon_start=1
/transl_table=11
/product="putative PTS system, enzyme IIB"
/protein_id="AAL98569.1"
/db_xref="GI:19749183"
/translacion="MIKIGLIFCAAGFSGTMLVNMKVAAEKGIDCQIEAYAQGKLAD
YAPLLDVALLGPOVAYTLDKSEAICKNDIPVIAIPMADYGMLDGNKVLDSLVRKE
"
/complement(7981..8301)
/gene="spym18_2114"
/complement(8301..8570)
/notes="best blastp match gb|AAK34717.1| (AB006626)
putative PTS system, enzyme III [Streptococcus pyogenes M1
GAS]"
/codon_start=1
/transl_table=11
/product="putative PTS system, enzyme III"
/protein_id="AAL98570.1"
/db_xref="GI:19749184"
/translacion="MOVIVPDQIIMGLINAGDAKOHVYQALCKAKEDDYATSEKWA
LADALLEAHNLQTFLAQASGNKSEITAFVHSQDHLMTTITEINLVKEIIDLRKE
LATK"
/complement(8590..9570)
/gene="spym18_2115"
/complement(9570..10424)
/notes="best blastp match gb|AAK34718.1| (AB006626)
putative transcriptional regulator [Streptococcus pyogenes
M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative transcriptional regulator"
/protein_id="AAL98571.1"
/db_xref="GI:19749185"
/translacion="MKEERRRLAKVAYLHYVQGSQTLISKEMNIYRTTVCRMLAKA
KEEGIVRIETADYADLFALREYVQOYGLKDLVPOVEDTPMDTLNVAKTAAEV
FRHVVDKDGKIGISWAGTISCLMDLNPAMKDVFIYPLAGGSPSHNAYHYNTLVYR
LARIHNSAFNMYIQEDKRLAKISQKFNIDILTSWDQDLALVIGIGEPNSLE
QSQRDLTSSDDHQKYEKAVEGVCRRFDQAGPVYTGLDQRTIGISLEQLRRVFK
"
Query Match 95.8%; Score 1037.2; DB 1; Length 10424;
Best Local Similarity 99.2%; Pred. No. 1.9e-22;
Matches 1042; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

QY

```

1 GACACGCCCTCTCTTTTCTCCTTAATTAATCTCGTTAATTTTCAATATTTTAAAAAAC 60
|||||

```

RESULT 8  
AB030578

```

Db
1050 GACACGCCCTCTCTTTTCTCCTTAATTAATCTCGTTAATTTTCAATATTTTAAAAAAC 991
QY
61 TATTGATAAACTAGTTAAAGCGTATACCTATGTTAGTTAGCGAAATTAGAAAAAGAG 120
|||||
990 TATTGATAAACTAGTTAAAGCGTATACCTATGTTAGTTAGCGAAATTAGAAAAAGAG 931
|||||
121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGCGGCTAGTAA 180
|||||
930 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGCGGCTAGTAA 871
|||||
181 AATTTTCAATGGTAGCTCTTGTATCAGCCCAATGGCTGTAAACAAAGTACACCTTCAAAA 240
|||||
870 AATTTTCAATGGTAGCTCTTGTATCAGCCCAATGGCTGTAAACAAAGTACACCTTCAAAA 811
|||||
241 ATACTGCACCTGGCAGCAACACACAGCTCTCAAAATGATGTTGTTCTAAATGATGGCCCAA 300
|||||
810 ATACTGCACCTGGCAGCAACACACAGCTCTCAAAATGATGTTGTTCTAAATGATGGCCCAA 751
|||||
301 GCAAGTACCTAAACGAAGCAATTAGCTTGGACATTCATGACAGCTCCTAACTATTACAAAA 360
|||||
750 GCAAGTACCTAAACGAAGCAATTAGCTTGGACATTCATGACAGCTCCTAACTATTACAAAA 691
|||||
361 CTTTAGTACTAGTCAAGTACTCCAGCAGCTCTTTCCTAAAGCAGGAGATATTCCTCTATA 420
|||||
690 CTTTAGTACTAGTCAAGTACTCCAGCAGCTCTTTCCTAAAGCAGGAGATATTCCTCTATA 631
|||||
421 GCAAAATTTAGATGATGTTAGGAAGGACGCGTACTGCTGAGAGTACATTTGACTTATGCCAATG 480
|||||
630 GCAAAATTTAGATGATGTTAGGAAGGACGCGTACTGCTGAGAGTACATTTGACTTATGCCAATG 571
|||||
481 TTGAAGGTAGCTACGGTGTGTAGACAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGGA 540
|||||
570 TTGAAGGTAGCTACGGTGTGTAGACAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGGA 511
|||||
541 CTGGAACCCCTAATCATGTCAAAATATAAATTCGAATGGTTAAATGGTCTATCTTATGTCG 600
|||||
510 CTGGAACCCCTAATCATGTCAAAATATAAATTCGAATGGTTAAATGGTCTATCTTATGTCG 451
|||||
601 GAGATTTCTGGAATAGAAATCATCTCATTTGACAGATAGTCTCGGTGGAGAGTGCATCAGAG 660
|||||
450 GAGATTTCTGGAATAGAAATCATCTCATTTGACAGATAGTCTCGGTGGAGAGTGCATCAGAG 391
|||||
661 TCAATGCCGTTACAGGAACACGTAGTACCAAAATGTAGAGAGTCTGTGACCAAAAAAGCGCGCA 720
|||||
390 TCAATGCCGTTACAGGAACACGTAGTACCAAAATGTAGAGAGTCTGTGACCAAAAAAGCGCGCA 331
|||||
721 TCGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780
|||||
330 TCGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 271
|||||
781 ATTATGAAGTCGCTCCAAATCTACACGACAGAGTGTGATTCCAAGAGCTGTGTGTTAT 840
|||||
270 ATTATGAAGTCGCTCCAAATCTACACGACAGAGTGTGATTCCAAGAGCTGTGTGTTAT 211
|||||
841 CAATGCAATCTTCTGATAATACCATCAACGAGAGAAAGTATTAGTTTACACACAGCTAATG 900
|||||
210 CAATGCAATCTTCTGATAATACCATCAACGAGAGAAAGTATTAGTTTACACACAGCTAATG 151
|||||
901 GCTACACCAATTAACCTACCATAACGGTACACCTACTCAAAAAATAATACCAAAAGGCTAGAC 960
|||||
150 GCTACACCAATTAATTAACCTACCATAACGGTACACCTACTCAAAAAATAATACCAAAAGGCTAGAC 91
|||||
961 CTCTGCTCTACTAGGCCCTAGCTTTTACATCAAAAAAGCAATGACTATATAGAAAAAGTAAAAA 1020
|||||
90 CTCTGCTCTACTAGGCCCTAGCTTTTACATCAAAAAAGCAATGACTATATAGAAAAAGTAAAAA 31
|||||
1021 TACTAGAAAAAGCAATGATTGCGGTCATTG 1050
|||||
30 TACTAGAAAAAGCAATGATTGCGGTCATTG 1
|||||

```



Db 1135 CTCTGCTCACTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGT-AAAA 1193

Qy 1021 TACTAGAAAAAGCAATGATCCGCGTCAATGCTTTTATGAATTTGTGCAAAAAGCAAA 1078  
|||||  
Db 1194 TACTAGAAAAAGCAATGATCCGCGTCAATGCTTTTATGAATTTGTGCAAAAAGTAAAA 1251  
|||||

RESULT 9

E08254

LOCUS DNA encoding mitogenic factor. 1021 bp DNA linear PAT 29-SEP-1997

ACCESSION E08254

VERSION E08254.1 GI:2176375

KEYWORDS JP 1994279498-A/1.

SOURCE Streptococcus pyogenes

ORGANISM Streptococcus pyogenes

REFERENCE 1 (bases 1 to 1021)

AUTHORS Yuzudou,T., Okumura,K., Iwasaki,M., Hara,A., Kishishita,M., Takeda,Y., Igarashi,H. and Hinuma,Y.

TITLE MITOGENIC FACTOR, ITS GENE AND METHOD FOR DETECTING THE SAME GENE

JOURNAL Patent: JP 1994279498-A 1 04-OCT-1994;

SHIONOGI & CO LTD

OS Streptococcus pyogenes

PN JP 1994279498-A/1

PD 04-OCT-1994

PF 28-JAN-1994 JP 1994026252

PR 01-FEB-1993 JP 93P 37383

PI YUZUDOU TAKASHI, OKUMURA KOUICHI, IWASAKI MAKOTO, HARA AYAKO, KISHISHITA MASAMICHI, TAKEDA YOSHIFUMI, IGARASHI HISANAGA, PI

HINUMA YORIO

PC C07K13/00,C12N15/31,C12P21/02,C12Q1/68,(C12N15/31,C12R1:465);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC Feature is identified by experimental;

FT Key Location/Qualifiers

FT source 1..1021 /organism='Streptococcus pyogenes' FT

FT /strain='NY-5',

FT mat\_peptide 155..967 /product='mitogenic factor'.

FT Location/Qualifiers

1..1021 /organism='Streptococcus pyogenes'

/mol\_type='genomic DNA'

/db\_xref='taxon:1314'

BASE COUNT 349 a 195 c 199 g 278 t

ORIGIN

Query Match 88.8%; Score 961.6; DB 6; Length 1021;

Best Local Similarity 98.5%; Pred. No. 2e-205;

Matches 981; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1 GACAACGCTCTCTTTTCTCCTTACTATCTCCTTTAAATTTTCAATTTTAAAAAAC 60  
|||||

Db 27 GACAATGCCCTCTTTTCTCCTTATTATCTCGTTTAAATTTTCAATTTTAAAAAAC 86  
|||||

Qy 61 TATTGATAAACTAGTTAAGTAAGCGGTATCTATGTTAGTGGTAAATAGAAAAGG 120

Db 87 TATTGATAAACTAGTTAAGTAAGCGGTATCTATGTTAGTGGTAAATAGAAAAGG 146

Qy 121 ACAAGCATATGAATCTACTTGGATCAAGCGGTTTTTCTAAAAATGCGGTAGTAA 180  
|||||

Db 147 ACAAGCATATGAATCTACTTGGATCAAGCGGTTTTTCTAAAAATGCGGTAGTAA 206  
|||||

Qy 181 AATTTTCAATGGTAGCTCTGTATCAGCCCAATGCGGTAAACACAGTCACACTTCAA 240  
|||||

Db 207 AATTTTCAATGGTAGCTCTGTATCAGCCCAATGCGGTAAACACAGTCACACTTCAA 266  
|||||

Qy 241 ATATGCACTGGCAGCAAAACACAGGTCTCAATGATGTTGTTCTAAATGATGGCGCAA 300  
Db 267 ATATGCACTGGCAGCAAAACACAGGTCTCAATGATGTTGTTCTAAATGATGGCGCAA 326  
Qy 301 GCAAGTACCTAAACGAAGCAATTAGCTTTGGACATTCATCAATGACAGTCCCTAACTATTACAAA 360  
|||||

Db 327 CGAAGTACCTAAACGAAGCAATTAGCTTTGGACATTCATCAATGACAGTCCCACTATTACAAA 386  
|||||

Qy 361 CTTTGTAGTACTAGTACAGATTAATCTCAGCACTCTTTTCTTAAAGCAGGAGATTTCTCTATA 420  
Db 387 CTTTGTAGTACTAGTACAGATTAATCTCAGCACTCTTTTCTTAAAGCAGGAGATTTCTCTATA 446  
Qy 421 GCAAAATTAGATGAGTTAGGAAGGCGGTACTGCTAGAGGTACATTGACATTATGCCAATG 480  
Db 447 GCAAAATTAGATGAGTTAGGAAGGCGGTACTGCTAGAGGTACATTGACATTATGCCAATG 506  
Qy 481 TTGAAGGTAGTACGCGTGTAGACAACTCTTTCGGTAAAAATCAAAACCCCGCAGATGGA 540  
Db 507 TTGAAGGTAGTACGCGTGTAGACAACTCTTTCGGTAAAAATCAAAACCCCGCAGATGGA 566  
Qy 541 CTGGAACCCCTAATCATGTCAAAATATAAAAATTGAATGGTTAAATGCTTATCTTATGTCG 600  
Db 567 CTGGAACCCCTAATCATGTCAAAATATAAAAATTGAATGGTTAAATGCTTATGTCG 626  
Qy 601 GAGATTTCTGGAATAGAAATCATCTCATTCAGATAGTCTCGGTGGAGATGCACCTCAGAG 660  
Db 627 GAGATTTCTGGAATAGAAATCATCTCATTCAGATAGTCTCGGTGGAGATGCACCTCAGAG 686  
Qy 661 TCAATGCGGTTACAGGAACACGTACCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCA 720  
Db 687 TCAATGCGGTTACAGGAACACGTACCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCA 746  
Qy 721 TGGCTATACCGAACAAAGAGCTCAAGAATGGTTAGGAAGCAAAATCGTGTGCTTATCTTT 780  
Db 747 TGGCTATACCGAACAAAGAGCTCAAGAATGGTTAGGAAGCAAAATCGTGTGCTTATCTTT 806  
Qy 781 ATTATGAAGTGGTCCAAATCTCAACGACAGGTTGATTCGAAGAGCTGTGCTGTTAT 840  
Db 807 ATTATGAAGCGGTCCAAATCTCAACGACAGCGAGTTGATTCGAAGAGCTGTGCTGTTAT 866  
Qy 841 CAATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAACACACTAATG 900  
Db 867 CAATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAACACACTAATG 926  
Qy 901 GCTACACCAATTAATCACTACCAATGCTACCTCTCAAAAATTAATACCAAAAGGCTAGAC 960  
Db 927 GCTACACCAATTAATCACTACCAATGCTACCTCTCAAAAATTAATACCAAAAGGCTAGAC 986  
Qy 961 CTCTGCTCACTAGCGCTAGCTTTTACATCAAAAAA 996  
Db 987 CTCTGCTCACTAGG-CTAGCTTTTACATCAAAAAA 1021

RESULT 10

A49205

LOCUS Sequence 11 from Patent WO9606174. 940 bp DNA linear PAT 07-MAR-1997

DEFINITION A49205

ACCESSION A49205

VERSION A49205.1 GI:2302736

KEYWORDS Streptococcus pyogenes

SOURCE Streptococcus pyogenes

ORGANISM Streptococcus pyogenes

REFERENCE 1 (bases 1 to 940)

AUTHORS Adams,C.W., Pang,P.P. and Belei,M.C.

TITLE RECOMBINANT DNase B DERIVED FROM STREPTOCOCCUS PYOGENES

JOURNAL Patent: WO 9606174-A 11 29-FEB-1996;

BECKMAN INSTRUMENTS INC (US)

FEATURES Location/Qualifiers

source 1..940 /organism='Streptococcus pyogenes'

/mol\_type='genomic DNA'



```
Db 721 ATGCAATCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGGC 780
Qy 903 TACACCAATTAACCTAACCGGTACACCTACTCAAAATAATACCAAAAGGCTAGACCT 962
Db 781 TACACCAATTAACCTAACCGGTACACCTACTCAAAATAATACCAAAAGGCTAGACCT 840
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAATA 1022
Db 841 CTGCTCACTAGGCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAATA 900
Qy 1023 CTAGAAAAGCAATGATTGCGCTCATTCG 1051
Db 901 CTAGAAAAGCAATGATTGCGCTCATTCG 929

RESULT 12
LOCUS A49208 937 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 14 from Patent WO9606174.
ACCESSION A49208
VERSION A49208.1 GI:2302739
KEYWORDS Streptococcus pyogenes
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 937)
AUTHORS Adams,C.W., Pang,P.P. and Belei,M.C.
TITLE RECOMBINANT Dnaase B DERIVED FROM STREPTOCOCCUS PYOGENES
JOURNAL Patent: WO 9606174-A 14 29-FEB-1996;
BECKMAN INSTRUMENTS INC (US)
FEATURES
    source
        1..937
            /organism="Streptococcus pyogenes"
            /mol_type="genomic DNA"
            /db_xref="taxon:1314"
            /notes="unnamed protein product"
            /codon_start=1
            /transl_table=11
            /protein_id="CAA03161.1"
            /db_xref="GI:2302740"
            /translation="MDPNLIGSRVFSKKRLVKFMSVALVSAATVTTVTLNTALA
            QTQVNDVLTNDGSKYLNALAWTNDSPNYKTLGTSQITPALFPKAGDILYSKLD
            ELGRTARTGLTYANVEGSGYGRQFGKQNPAGWTPGNPHVKYKIEWNLGSLVYGD
            FMRSHLIADSLGDLARVNAVTRTQNVGGRDQKGMRYTEQRAQEWLEARNRQGYL
            YVEVAPLYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGTYINHNGTPTOK"

BASE COUNT 301 a 203 c 192 g 241 t
ORIGIN
    Query Match 78.8%; Score 853.8; DB 6; Length 937;
    Best Local Similarity 95.7%; Pred. No. 3.1e-181;
    Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATGTTCGGCTAGTAAA 182
Db 1 ATGGATCCGAACCTGCTGGGTTCCCGTCGTGTTTCTCCAAAAAATGCGCTGTGTTAA 60
Qy 183 TTTTCAATGTAGCTCTTGATCAGCCACAATGGCTGTACACAGCTCACACTTGAAT 242
Db 61 TTCTCCATGGTTGCTCTGGTTTCGGCTACCATGGCTGTACCACCGGTACCTGGAAAC 120
Qy 243 ACTGCACTGGCAGCAGCAAAACACAGGCTCTCAAAATGATGTTCTTAATGATGCGCAAGC 302
Db 121 ACCGCTCTGCG---TCAGACACAGGCTCAAAATGATGTTGTTCTTAATGATGCGCAAGC 177
Qy 303 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGCTCCTAACTATTACAAACT 362
Db 178 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGCTCCTAACTATTACAAACT 237
Qy 363 TTAGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTCTCTATAGC 422
```

```
Db 238 TTAGGTACTAGTCAGATTACTCCAGCACCTCTTTCTTAAAGCAGGAGATATTCTCTATAGC 297
Qy 423 AAATTAGATGAGTTAGGAAGGACGCGTACTGCTAGAGGTACATTTGCTTATGCGCAATGTT 482
Db 298 AAATTAGATGAGTTAGGAAGGACGCGTACTGCTAGAGGTACATTTGCTTATGCGCAATGTT 357
Qy 483 GAAGGTAGCTACCGTGTAGACAACTCTTTCCGGTAAATAATCAAAACCCCGCAGGATGACT 542
Db 358 GAAGGTAGCTACCGTGTAGACAACTCTTTCCGGTAAATAATCAAAACCCCGCAGGATGACT 417
Qy 543 GGAACCCCTAATCATGTCAAAATATAAAATGTTAAATGGTTAAATGGTCTTATGTGCGGA 602
Db 418 GGAACCCCTAATCATGTCAAAATATAAAATGTTAAATGGTCTTATGTGCGGA 477
Qy 603 GATTTCTGGAATAGAAGTCAATCATTTGCGATAGTCTCGGTGGAGATGCACTCAGAGTC 662
Db 478 GATTTCTGGAATAGAAGTCAATCATTTGCGATAGTCTCGGTGGAGATGCACTCAGAGTC 537
Qy 663 AATGCCGTTACAGGAACACGTACCCAAATGTAGGAGGTGCTGACCAAAAGGCGCATG 722
Db 538 AATGCCGTTACAGGAACACGTACCCAAATGTAGGAGGTGCTGACCAAAAGGCGCATG 597
Qy 723 CGCTATACCGAACAAAGAGCTCAAGAAATGGTTAGAAGCAAAATCGTATGGCTATCTTTAT 782
Db 598 CGCTATACCGAACAAAGAGCTCAAGAAATGGTTAGAAGCAAAATCGTATGGCTATCTTTAT 657
Qy 783 TATGAAGTCGCTCCAAATCTACAAACGAGAGCTTCAATCCAAAGAGCTGTCGTTGATCA 842
Db 658 TATGAAGTCGCTCCAAATCTACAAACGAGAGCTTCAATCCAAAGAGCTGTCGTTGATCA 717
Qy 843 ATGCAATCTCTGATATACCATCAACGAGAGAAAGTATTAGTTTACACACAGCTAATGGC 902
Db 718 ATGCAATCTCTGATATACCATCAACGAGAGAAAGTATTAGTTTACACACAGCTAATGGC 777
Qy 903 TACACCATTAACCTACCATAACGCTACACCTACTCAAAATAATATACCAAAAGGCTAGACCT 962
Db 778 TACACCATTAACCTACCATAACGCTACACCTACTCAAAATAATATACCAAAAGGCTAGACCT 837
Qy 963 CTGCTCACTAGGCTAGCTATTTTACATCAAAAGCAATGACTATAGAAAGTAAATA 1022
Db 838 CTGCTCACTAGGCTAGCTATTTTACATCAAAAGCAATGACTATAGAAAGTAAATA 897
Qy 1023 CTAGAAAAGCAATGATTGCGCTCATTCG 1051
Db 898 CTAGAAAAGCAATGATTGCGCTCATTCG 926

RESULT 13
AR219229
LOCUS AR219229 937 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 14 from patent US 6420152.
ACCESSION AR219229
VERSION AR219229.1 GI:23320177
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 937)
AUTHORS Adams,C.W., Pang,P.P.Y. and Belei,C.M.
TITLE Recombinant Dnaase B derived from Streptococcus pyogenes
JOURNAL Patent: US 6420152-A 14 16-JUL-2002;
FEATURES
    Location/Qualifiers
        source
            1..937
                /organism="unknown"
                /organism 301 a 203 c 192 g 241 t
                BASE COUNT
                ORIGIN

Query Match 78.8%; Score 853.8; DB 6; Length 937;
Best Local Similarity 95.7%; Pred. No. 3.1e-181;
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATGTTCGGCTAGTAAA 182
```

```
Db 1 ATGATCCGAACCTGCTGGGTTCCCGTCGTGTTTCTCCAAAAATGCCGCTCGTTTAAA 60
Qy TTTTCATGTAGTCTCTGTATCAGCCAAATGGCTGTAAACACAGTCACACTCAAAAT 242
Db TTTTCATGTAGTCTCTGTATCAGCCAAATGGCTGTAAACACAGTCACACTCAAAAT 120
Qy ACTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302
Db ACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 177
Qy AAGTACTTAAACGAAGCATAGCTGGACATTCATCAATGACAGTCCTTAACTTACAAACT 362
Db AAGTACTTAAACGAAGCATAGCTGGACATTCATCAATGACAGTCCTTAACTTACAAACT 237
Qy TTAGTACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 422
Db TTAGTACTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 297
Qy AAATTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 482
Db AAATTAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
Qy GAAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
Db GAAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417
Qy GAAACCTTAAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 602
Db GAAACCTTAAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 477
Qy GATTTCTGGAATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 662
Db GATTTCTGGAATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537
Qy AATGCGTTTACAGAAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722
Db AATGCGTTTACAGAAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
Qy CGCTATACCGAAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
Db CGCTATACCGAAACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657
Qy TATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 842
Db TATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
Qy ATGCAATCTTCTGATTAATACCATCAACGAGAGTGTATGTTTACCAACAGTCTGCTGCTGCTGCT 902
Db ATGCAATCTTCTGATTAATACCATCAACGAGAGTGTATGTTTACCAACAGTCTGCTGCTGCTGCT 777
Qy TACACATTAATACCATCAACGAGAGTGTATGTTTACCAACAGTCTGCTGCTGCTGCTGCTGCT 962
Db TACACATTAATACCATCAACGAGAGTGTATGTTTACCAACAGTCTGCTGCTGCTGCTGCTGCT 837
Qy CTGCTACCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1022
Db CTGCTACCTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
Qy CTAGAAAAGCAATGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
Db CTAGAAAAGCAATGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926

RESULT 14
SPV295272
LOCUS Streptococcus pyogenes dnaseB gene for streptodornase B.
DEFINITION Streptococcus pyogenes dnaseB gene for streptodornase B.
ACCESSION AJ295272
VERSION AJ295272.1 GI:13539552
KEYWORDS dnaseB gene; streptodornase B.
SOURCE Streptococcus pyogenes
ORGANISM Streptococcus pyogenes
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
```

```
Streptococcus.
REFERENCE
1 Gerlach, D., Schmidt, K.H. and Fleischer, B.
AUTHORS Basic streptococcal superantigens (SPEX/SMEZ or SPEC) are
TITLE responsible for the mitogenic activity of the so-called mitogenic
factor (MF)
JOURNAL FEMS Immunol. Med. Microbiol. 30 (3), 209-216 (2001)
MEDLINE 21233096
PUBMED 11335140
REFERENCE
2 (bases 1 to 816)
AUTHORS Schmidt, K.H.
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2000) Schmidt K.H., Institute of Medical
Microbiology, Friedrich-Schiller-University, Semmelweisstrasse 4,
D-07740, GERMANY
FEATURES
Location/Qualifiers
1..816
/organism="Streptococcus pyogenes"
/mol_type="genomic DNA"
/strain="C2038"
/db_xref="taxon:1314"
1..816
/gene="dnaseB"
1..816
/gene="dnaseB"
/function="cleavage of DNA"
/codon_start=1
/transl_table=11
/product="streptodornase B"
/protein_id="CAC35734.1"
/db_xref="GI:13539553"
/translation="MNLIGSRVFSKCRVLFKSMVALVSATMVTVTLENTALAQ
TOVNDVNLDCASKYLNEALAWTENDSPNYKLTQTSQITPALFPKAGDILYSKJDE
LGRKTARTGLTYANVEGSGYGRQSPKQNPAGWPNHVKYKIEWLNLGSLYGVDF
WNRSLIADSLGGDALRVNAVITGRTQNVGDRDQKGMRYTEQRAQWLEARNRGYLY
YEVAIPYNADLEIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINHNGTPTQK"
sig_peptide 1..126
mat_peptide 127..813
BASE COUNT 276 a 162 c 166 g 212 t
ORIGIN
Query Match 75.3%; Score 816; DB 1; Length 816;
Best Local Similarity 100.0%; Pred. No. 9.5e-173;
Matches 816; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 129 ATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAAATTTTCA 188
Db 1 ATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAAATTTTCA 60
Qy 189 ATGGTAGCTCTTGTATCAGCCCAATGGCTGTAAACACAGTCACACTTGAANAATCTGCA 248
Db 61 ATGGTAGCTCTTGTATCAGCCCAATGGCTGTAAACACAGTCACACTTGAANAATCTGCA 120
Qy 249 CTGCGACACACAAACACAGAGTCTCAAAATGATGTTGTTCTTAAATGATGCGCAAGCAAGTAC 308
Db 121 CTGCGACACACAAACACAGAGTCTCAAAATGATGTTGTTCTTAAATGATGCGCAAGTAC 180
Qy 309 CTAAACGAAGCATAGCTTGGACATTCATCAATGACAGTCCTTAACTATTACAAAATCTTAGGT 368
Db 181 CTAAACGAAGCATAGCTTGGACATTCATCAATGACAGTCCTTAACTATTACAAAATCTTAGGT 240
Qy 369 ACTAGTCAGATTACTCCAGACTCTTTCTTAAAGCAGGAGATATTCTCTATAGCAAAATTA 428
Db 241 ACTAGTCAGATTACTCCAGACTCTTTCTTAAAGCAGGAGATATTCTCTATAGCAAAATTA 300
Qy 429 GATGAGTTAGGAGGACGGGTACTGCTAGAGGTACATTGACTTATGCGCAATGTTGAAGGT 488
Db 301 GATGAGTTAGGAGGACGGGTACTGCTAGAGGTACATTGACTTATGCGCAATGTTGAAGGT 360
Qy 489 AGCTACGGTGTAGACAATCTTTCGGTAAAAAATCAAAACCCCGCAGGATGAGTGGAAAC 548
```

Db	361	AGCTACGGTGTAGACAATCTTTTCGGTAAAAAATCAAAACCCCGCAGATGCACTGGAAAC	420
Qy	549	CCTAATCATGTCAAAATATATAAATTGAATGGTAAATGGTCTATCTTATGTCGGAGATTTTC	608
Db	421	CCTAATCATGTCAAAATATATAAATTGAATGGTAAATGGTCTATCTTATGTCGGAGATTTTC	480
Qy	609	TGGAATAGAAGTCATCTCATTCGAGATAGTCTCGGTGGAGATGCACCTACAGAGTCAATGCC	668
Db	481	TGGAATAGAAGTCATCTCATTCGAGATAGTCTCGGTGGAGATGCACCTACAGAGTCAATGCC	540
Qy	669	GTTACAGGAAACGTCATCCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCATGCGCTAT	728
Db	541	GTTACAGGAAACGTCATCCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCATGCGCTAT	600
Qy	729	ACCGAAACAAAGAGCTCAAGAAATGGTTAGAACAATCGTGATGGCTATCTTTATATATGAA	788
Db	601	ACCGAAACAAAGAGCTCAAGAAATGGTTAGAACAATCGTGATGGCTATCTTTATATATGAA	660
Qy	789	GTCGCTCCAATCTCAACGCGACAGAGTTGATTTCAAGAGCTGTGCTGGTATCAATGCCAA	848
Db	661	GTCGCTCCAATCTCAACGCGACAGAGTTGATTTCAAGAGCTGTGCTGGTATCAATGCCAA	720
Qy	849	TCTTCTGATATACCATCAACGAGAAAGTATTAGTTTCAACACAGCTTAATGGCTACACC	908
Db	721	TCTTCTGATATACCATCAACGAGAAAGTATTAGTTTCAACACAGCTTAATGGCTACACC	780
Qy	909	ATTAACTACATACGCGTACACCTACTCAAAAATAA	944
Db	781	ATTAACTACATACGCGTACACCTACTCAAAAATAA	816
RESULT 15	AE014158/c		
LOCUS	AE014158	5081 bp	DNA linear BCT 18-JUL-2002
DEFINITION	Streptococcus pyogenes MGAS315, section 23 of the complete genome.		
ACCESSION	AE014158	AE014074	
VERSION	AE014158.1	GI:21904816	
KEYWORDS	Streptococcus pyogenes MGAS315		
SOURCE	Streptococcus pyogenes MGAS315		
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.		
REFERENCE	1 (bases 1 to 5081)		
AUTHORS	Beres,S.B., Sylva,G.L., Barbican,K.D., Lei,B., Hoff,J.S., Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F., Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and Musser,J.M.		
TITLE	Genome sequence of a serotype M3 strain of group A Streptococcus: Phenage-encoded toxins, the high-virulence phenotype, and clone emergence		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (15), 10078-10083 (2002)		
REFERENCE	2 (bases 1 to 5081)		
AUTHORS	Beres,S.B., Sylva,G.L., Barbican,K.D., Lei,B., Hoff,J.S., Mammarella,N.D., Liu,M.-Y., Smoot,J.C., Porcella,S.F., Parkins,L.D., McCormick,J.K., Leung,D.Y.M., Schlievert,P.M. and Musser,J.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-JUN-2002) Laboratory of Human-Bacterial Pathogenesis/ Rocky Mountain Laboratories, NIAID, NIH, 903 South Fourth St., Hamilton, MT 59840, USA		
FEATURES	Location/Qualifiers		
source	join(1..15324,49744..5081)		
	/organism="Streptococcus pyogenes MGAS315"		
	/mol_type="genomic DNA"		
	/strain="MGAS315"		
	/serotype="M3"		
	/db_xref="taxon:198466"		
	/focus		
	/note="group: A"		
source	complement(15325..49743)		
	/organism="Streptococcus pyogenes phage 315.3"		



```
/translation="MDFFAIDPIGLPHISLISLYLLSLLIALLLVLTFOAYRLKSYRY
FFLFLQSLQGLVTVYVLFDFLDELPHICRIAMTAIFLPRNRFKQLFVWLGI
GTFLLSLSDPLVPRLMHNVFVYGHVALLVNGLIYLLRFYDASQLRLSVRYL
ATVNFLLLSLAKGNIGFVMDIPVITHRLHLENFVITVITSLTEWVKITEFYLLKFG
EAQLAFASKEK"
gene      complement(4611..5978)
/genes="murF"
/notes="synonym: SpyM3_1083"
CDS       complement(4611..5978)
/genes="murF"
/notes="best blastp hit: gb|AAK34232.1| (AE006578) putative
D-Ala-D-Ala adding enzyme [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative D-Ala-D-Ala adding enzyme"
/protein_id="AAW79690.1"
/db_xref="GI:21904821"
/translation="MKLILHEVAKIVDAQNNVSDLDVPLHHIFDFSRKITKGLDPLP
LKGQRNGHEFIDLAFQNGAVATFSEKELPGKPHLLVEDCLKAFOKLAHYVIDKMRVDP
IATVSGNKTSTKMGIAVLTSTYKTYKQGNVNEIGLPTVILHMPDDPEKIVLEMG
QHMGIIRLLSEIARPIAVLTVEAHLEYFGSRDKIAQCKMQLVDGMDNGILLIAP
GSIIDPVLPEQNVIRFEDHQELNVTDIREEKDLSFTTNVLATPVSLPLPKYKYNAT
NAMVAAYVGLKILAVADEIIIEALESVOLTNRTWKKAANGADILSDVYNANPARNL
ILETANAKNPGGKKIAVLADMKELGPDVALHSQILKQLSPDRIDQLVFFVGDDIQE
LAQLAEVPEGHVQFQKNQOEQFEDMVRVQOQLLAPSDQILLKGSNSMGLAKLVN
HLVEN"
gene      complement(6177..7142)
/genes="ddlA"
/notes="synonym: SpyM3_1084"
CDS       complement(6177..7142)
/genes="ddlA"
/notes="best blastp hit: gb|AAK34233.1| (AE006579) putative
D-Ala-D-Ala ligase [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative D-Ala-D-Ala ligase"
/protein_id="AAW79691.1"
/db_xref="GI:21904822"
/translation="MRVNVYDKPLVTVITOMGOFIKTOQSPSSSRRLMNETI
ELTKIKPSDIYEGAVRPLVLPNGEDGSIQGLEFLVRLPMTYIGTNVMSSSIAMDKI
TKRVLSEIGIPQVATVYLDGDLAECLVETLARLTFFPVKPNMGSSVGIKAOT
KVELRAIQALVDSRVLEQSVAREIEVGLGNDKVKSTLPGEVTKDVFVDYQA
TYVDNKITMAIPADVQSVITMRSYAEVAFKALGGCGLSRCDPLFTQDGVLYNELN
TMPGFQMSYFILLWENMGLAYPDLIELVTLVLAIFQDRESHLI"
gene      complement(7324..7920)
/genes="recR"
/notes="synonym: SpyM3_1085"
CDS       complement(7324..7920)
/genes="recR"
/notes="best blastp hit: sp|Q99233|RECR_STRPY Recombination
protein recR"
/codon_start=1
/transl_table=11
/product="recombination protein recR"
/protein_id="AAW79692.1"
/db_xref="GI:21904823"
/translation="MLYTPAKLIDSKLPGGIKATRLAFYTGMSNEDVNDFA
KNLLAAKRELTVCSGNLMDDPCHICTDSRQDTILVVEDAKDVSAWEKIQEVHG
YVHVHLGISPMNGVPPDINLSLITRLMDGKVSIVVATNATADGEATSMYISRLV
KPAQIKVRLRGLAVGSDIEYADEVTLRAIENRTSL"
gene      complement(7967..8158)
/genes="SpyM3_1086"
CDS       complement(7967..8158)
/genes="SpyM3_1086"
/notes="best blastp hit: gb|AAC44614.1| (U58210)
penicillin-binding protein 2b [Streptococcus
thermophilus]"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAW79693.1"
/db_xref="GI:21904824"
/translation="MANPNSVVTIHKIGTAEITCITDNSDQTLVTYNLNVVAYNTNNE
```

```
DNIAVAVMYPHITNISAKM"
gene      complement(8719..9498)
/genes="fdhC"
/notes="synonym: SpyM3_1087"
CDS       complement(8719..9498)
/genes="fdhC"
/notes="best blastp hit: gb|AAK34235.1| (AE006579) putative
formate dehydrogenase [Streptococcus pyogenes M1 GAS]"
/codon_start=1
/transl_table=11
/product="putative formate dehydrogenase"
/protein_id="AAW79694.1"
/db_xref="GI:21904825"
/translation="MKTPEQILEATHIGHEHVKTKFLAKSILGIFGGMISIGYLLY
VRIASGLEAFGAPFISVIGACAPFIIILIMAGGELLITGNMMAVSAALLAKKIFSE
Query Match      25.6%; Score 276.8; DB 1; Length 50081;
Best Local Similarity 58.9%; Pred. No. 8.4e-52;
Matches 567; Conservative 0; Mismatches 362; Indels 33; Gaps 4;
QY 31 TCCTTTAATTTTCATATTTTAAAAAATTTTAACTATTTGATAAATAAGTAAAGCGGTATAC 90
Db 16659 TCGGTTAAATTTCAAAATGTTTAAAAAAGATTTGATTAAACCAATTAAGTAAACATATAA 16600
QY 91 TATGGTTAGTTAGCGAAATTTAGAAAAAGAGACAGCATATGAATCTTACTTGGATCAAGAC 150
Db 16599 TGTAAATTAGTGAGATAAAAAATAAAAAAGAGGTCGAAACATGAAC----- 16552
QY 151 GGGTTTTTTCTAAAAATGTTCGGCTAGTAAATTTTCAATGGTAGCTCTTGTATCAGCA 210
Db 16551 -----AACATAAGATTACCGTAAATTTTATTCTAGCATATTTCTAATCTTCTTAA 16503
QY 211 CAATGCTGTAAACACAGTCACACTTGAAATATCTGCACTGGCAGCAGCAAAACACAGGTC 270
Db 16502 TATCATCTCTATATCATTTATCATATCAACAGGCTACTTTTACG---TGATACTAAGAAG 16446
QY 271 CAATCATCTCTTCTTAATGATGCGGCAAGCAAGTACTTAAACGAAGCATTTAGCTTGA 330
Db 16445 TTAATGTTCTTAATAGATATAACTCTCAGATACATATCTACCTGAAAGCTTTTGTCTGGA 16386
QY 331 CATTCATGACAGTCTTAACTATTACAAACTTTTAGTACTAGTCAGATCTACTCCAGAC 390
Db 16385 CATTAGAAACTTCCCAAAATTTATTATAAGTATTAGGAGAAAGTGAATAGTAGAAAATT 16326
QY 391 TCTTTCTTAAACGAGGAGATATCTCTATAGCAATTTAGATGATTTAGGAGGACGCGTA 450
Db 16325 TATTTCCACAAAAGGCGCAATAGTCTATGCTGTCTTGACAGTCTTGGCAGAACATTAA 16266
QY 451 CTGCTAGAGGTACATTGACCTTATGCCAATGTTGAAGGTAGCTAGCGTGTGTAGACAATCTT 510
Db 16265 CAGTTAGAGGAACCTTGCACCTTCAATATGTTGAGTGTGTACATATAAGAAAAGATT 16206
QY 511 T-----CGGTAAAAATCAAAACCCCGAGGTGGACTGGAAACCCCTAATCATGTCAAAAT 564
Db 16205 TTAACGTTAGTAAGCAGAAACATTTGTGAGTGGCTAGGAAATAAATAATGCGGAAGTCT 16146
QY 565 ATAAAAATTGAATGGTTAAATGGTCTATCTTATCTCCGAGATTTCTGGAATAGAGTCAATC 624
Db 16145 ATGTTATAAAGGGTTAGGTGATGATTTCTATCAAGGTTATTTTGGGAATAGAGTCATT 16086
QY 625 TCATTGCGAGATAGTCTCGGTGGAGATGCACTCAGAGTCAATGCCGTTACAGGAACACGTA 684
Db 16085 TAATTGCGAGATAGTCTAGGCGGAGATGCCCTTAGAGTAAATGCTATTACAGGAACAGAA 16026
QY 685 CCCAAAATGTAGAGGTGCGTACCAAAAAGGCGGATCGGCTATATACCGAAACAAAGAGCTC 744
Db 16025 CACAAAATGTGAGGTGCTTCTGCAATGAGGATGAGATATACAGAAAATAAATCTC 15966
QY 745 AGAATGGTTAGAGCAATCTGTATGCTATCTTTATTATGAAGTCCCTCCAATCTACA 804
Db 15965 AGAATGGTTAGAGCCCATCGTGACGGTTACCTTTTACTATGAAGCAATGCCAATATACC 15906
QY 805 ACCGACGAGGTGATTCCAGAGCTGCTGCTGTTGATTCAATGCAATCTTCTGATTAATACCA 864
```





**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:53:48 ; Search time 348 Seconds  
(without alignments)  
8400.837 Million cell updates/sec

Title: US-08-482-785-7

Perfect score: 1083

Sequence: 1 GACAACGCCTCTCTTTTCT.....TGTCAAAGCAAAAGC 1083

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1079.8	99.7	1083	17	AAT12774 S. pyogenes DNaseB
2	958.4	88.5	1021	15	AAQ71612 Mitogenic factor g
3	921	85.0	944	17	AAT12780 DNase B modified g
4	919.4	84.9	944	16	AAQ85037 DNA-ase-B gene for
5	798.6	73.7	813	24	ABN70548 Streptococcus poly
6	400	36.9	400	16	AAQ85036 DNA-ase-B2 gene fr
7	400	36.9	400	17	AAT12773 S. pyogenes DNaseB
8	200	18.5	200	16	AAQ85040 DNA-ase-B gene ups

9	106.6	9.8	182	17	AAT12778 DNase B PCR primer
10	88.4	8.2	966	21	AAK99137 966nt DNA sequence
11	47.2	4.4	2409	24	ABN91441 Staphylococcus epi
C 12	46.2	4.3	6106	22	AA546429 Tumour suppressor
C 13	46.2	4.3	6106	24	ABK40031 Human chemically p
C 14	46.2	4.3	6106	24	ABL33472 Human immune syste
C 15	45	4.2	126	24	ABN70213 Streptococcus poly
16	44.6	4.1	2205	21	AA64207 Complete sequence
17	44	4.1	2406	22	AAS64207 S. epidermidis ope
C 18	44	4.1	2942	22	AH55066 S. epidermidis gen
C 19	44	4.1	2976	22	AH54836 S. epidermidis gen
C 20	44	4.1	3350	22	AH54295 S. epidermidis gen
21	44	4.1	6013	24	AAS61264 Human gene regulat
22	44	4.1	6013	24	ABK31360 Signal transductio
C 23	44	4.1	11178	24	ABL70508 Chemically treated
C 24	44	4.1	11178	24	AAS61060 Human gene regulat
C 25	44	4.1	11178	24	ABK31173 Signal transductio
C 26	43.4	4.0	8588	22	AA545469 Chemically pretrea
C 27	43.4	4.0	8588	24	ABK28325 DNA transcription
C 28	43.4	4.0	16373	24	ABL32618 Human immune syste
C 29	43.4	4.0	16373	24	AAD28382 Human chemically t
C 30	43.2	4.0	9770	24	ABL32032 Human immune syste
C 31	43.2	4.0	116277	20	AAK20249 Borrelia burgdorfe
C 32	43.2	4.0	910715	20	AAK20248 Human immune syste
C 33	42.6	3.9	13449	24	ABL33384 Human immune syste
C 34	42.4	3.9	7503	24	ABQ67009 Human angiogenesis
C 35	42.4	3.9	7503	24	ABL33548 Human immune syste
C 36	41.6	3.8	4590	22	AAH24065 Yeast AOD9604-asso
37	41.4	3.8	3344	22	AAH15978 Human cDNA sequenc
38	41.4	3.8	3484	21	AAK81047 Human secreted pro
C 39	41.4	3.8	5413	22	AA546693 Tumour suppressor
C 40	41.2	3.8	37184	24	ABQ67078 Human angiogenesis
C 41	41	3.8	5678	24	ABL33139 Human immune syste
C 42	40.8	3.8	1049	22	ABF97940 Human secreted pro
C 43	40.8	3.8	6101	24	ABL33474 Human immune syste
C 44	40.8	3.8	8238	24	ABL33988 Human immune syste
C 45	40.8	3.8	8238	24	AAS63348 Chemically pretrea

#### ALIGNMENTS

#### RESULT 1

AAT12774

ID AAT12774 standard; DNA; 1083 BP.

XX

AC AAT12774;

XX

DT 25-JUN-1996 (first entry)

XX

DE S. pyogenes DNaseB and leader sequence DNA.

XX

XX DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;

KW diagnosis; Escherichia coli; ds.

XX

OS Streptococcus pyogenes strain ATCC 14289.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



XX	Claim 1; Page 12-13; 20pp; English.
PS	
XX	
CC	AAQ17612 encodes a mitogenic factor which exhibits rabbit peripheral
CC	blood lymphocyte mitogenicity and/or DNA hydrolysing activity. It is
CC	strongly associated with group A Streptococci and the nucleotide
CC	sequences can be used for the microdetection of the gene and provide
CC	an early diagnosis of infectious disease caused by the bacteria.
CC	(See also AAQ17613-26).
CC	(Updated on 25-MAR-2003 to correct PN field.)
XX	
SO	Sequence 1021 BP; 349 A; 197 C; 197 G; 278 T; 0 other;
	Query Match            88.5%; Score 958.4; DB 15; Length 1021;
	Best Local Similarity 98.3%; Pred. No. 9.3e-239;
	Matches 979; Conservative 0; Mismatches 16; Indels 1; Gaps 1;
Qy	1 GACAAAGCCTTCTTTTTCCTTACTATCTPCCTTTAATTTCATATTTTTTAAAAAAC 60
Db	
Dy	27 GACANTGCCITCTTTTTTCTCCTTATTATCTCGTTAATTTTCATATTTTTTAAAAAAC 86
Qy	61 TATTGATAAACTAGTTAAGTAAGCGTATACATGCTTAGTTAGCGAAATAGAAAAGAGG 120
Db	
Dy	87 TAITTGATAAACTAGTTAAGTAAGCGTATACATGCTTAGTTAGCGAAATAGAAAAGAGG 146
Qy	121 ACAAGCATATGAATCTACTTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
Db	
Dy	147 ACAAGCATATGAATCTACTTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTGA 206
Qy	181 AATTTTCAAATGGTAGCTCTTGATCAGCCACAATGGGTGTAAACAACAGTACACTTCAAA 240
Db	
Dy	207 AATTTTCAAATGGTAGCTCTTGATCAGCCACAATGGGTGTAAACAACAGTACACTTGA 266
Qy	241 ATACTGCATGGGCACGACAAACACAGGTCTCAAAATGATGTGTTCTAAATGATGCGCGAA 300
Db	
Dy	267 ATACTGCATGGGCACGACAAACACAGGTCTCAAAATGATGTGTTCTAAATGATGCGCGAA 326
Qy	301 GCAAGTACCTTAACGAGACATTAGCTTGGACATTCATGACAGTCCTAACATATTACAAA 360
Db	
Dy	327 CGAAGTACCTTAACGAGACATTAGCTTGGACATTCATGACAGTCCCAACTATTACAAA 386
Qy	361 CTTTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATCTCTATA 420
Db	
Dy	387 CCTTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATCTCTATA 446
Qy	421 GC AAA TTAGATGAGTTAGSAGGACGGTACTGCTAGAGGTACATTCGATATGCCCCAATG 480
Db	
Dy	447 GC AAA TTAGATGAGTTAGSAGGACGGTACTGCTAGAGGTACATTCGATATGCCCCAATG 506
Qy	481 TTGAAGTACTACGGTGTAGACAATCTTTTCGGTAAATTCAAAACCCCGCAGGATGGA 540
Db	
Dy	507 TTGAAGTACTACGGTGTAGACAATCTTTTCGGTAAATTCAAAACCCCGCAGGATGGA 566
Qy	541 CTGGAAACCCTAATCATGTCAAATATAAAAATTCGAATGGTTAAATGGTCTATCTTATGTCG 600
Db	
Dy	567 CTGGAAACCCTAATCATGTCAAATATAAAAATTCGAATGGTTAAATGGTCTCTTATGTCG 626
Qy	601 GAGATTTCTGGAATAGAAAGTCATCTCATTTGCGAGATGCTCCGTTGGAGATGCACTCAGAG 660
Db	
Dy	627 GAGATTTCTGGAATAGAAAGTCATCTCATTTGCGAGATGCTCCGTTGGAGATGCACTCAGAG 686
Qy	661 TCAATGCGGTTACAGGAACACGTACCCAAAATGTPAGGAGTTCGTGACCAAAAAGGCGGCA 720
Db	
Dy	687 TCAATGCGGTTACAGGAACACGTACCCAAAATGTPAGGAGTTCGTGACCAAAAAGGCGGCA 746
Qy	721 TGCCTATACCGAACAAAGAGCTCAAGAATGGTTAGAAAGCAAAATCGTATGGCTATCTTT 780
Db	
Dy	747 TGCCTATACCGAACAAAGAGCTCAAGAATGGTTAGAAAGCAAAATCGTATGGCTATCTTT 806
Qy	781 ATTATGAAGTCGCTCCAATCTCAACCCAGACAGTTGATTCNAGAGCTGTCGTGGTAT 840
Db	
Dy	807 ATTATGAAGTCGCTCCAATCTCAACCCAGACAGTTGATTCNAGAGCTGTCGTGGTAT 866

```
QY 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAATGTCGGCTAGTAAA 182
DB 1 ATGGATCCGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAATGTCGGCTAGTAAA 60
QY 183 TTTTCAATGTAGTCTTGTATCAGCCCAATGGCTGTAAACAGTCACTTGAAT 242
DB 61 TTTTCAATGTAGTCTTGTATCAGCCCAATGGCTGTAAACAGTCACTTGAAT 120
QY 243 ACTGCATGCGCAGCAACACACAGCTCTCAATGATGTTGTTCTAAATGATCGCGCAAGC 302
DB 121 ACTGCATGCGCAGCAACACACAGCTCTCAATGATGTTGTTCTAAATGATCGCGCAAGC 180
QY 303 AAGTACTTAAACGAAGCATTTAGCTTGGACATTCATCAATGACAGCTCTCAATGATCAAACT 362
DB 181 AAGTACTTAAACGAAGCATTTAGCTTGGACATTCATCAATGACAGCTCTCAATGATCAAACT 240
QY 363 TTAGGTACTAGTCCAGATTAAGTCTCCAGCACTCTTCTTAAAGAGGAGATTTCTTATAGC 422
DB 241 TTAGGTACTAGTCCAGATTAAGTCTCCAGCACTCTTCTTAAAGAGGAGATTTCTTATAGC 300
QY 423 AAATTAGATGAGTTAGGAAGGACGCTACTGCTAGAGGTACATTCGACTTATGCCAATGTT 482
DB 301 AAATTAGATGAGTTAGGAAGGACGCTACTGCTAGAGGTACATTCGACTTATGCCAATGTT 360
QY 483 GAAGGTAGTACGGTGTAGACAAATCTTTTCGGTAAATAATCAAAACCCCGCAGGATGACT 542
DB 361 GAAGGTAGTACGGTGTAGACAAATCTTTTCGGTAAATAATCAAAACCCCGCAGGATGACT 420
QY 543 GGAACCCCTTAATCATGTCAATATATAAATTGAATGGTTAAATGGTCTATCTTATGCGGA 602
DB 421 GGAACCCCTTAATCATGTCAATATATAAATTGAATGGTTAAATGGTCTATCTTATGCGGA 480
QY 603 GATTTCTGGAATAGAGTCAATCTCAATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 662
DB 481 GATTTCTGGAATAGAGTCAATCTCAATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 540
QY 663 AATGCGGTTTACAGGAACAGTACCCAAATGTAGAGGTCTGTGACCAAAAAGGCGGATG 722
DB 541 AATGCGGTTTACAGGAACAGTACCCAAATGTAGAGGTCTGTGACCAAAAAGGCGGATG 600
QY 723 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAACAAATCGTGTGCTATCTTTAT 782
DB 601 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAACAAATCGTGTGCTATCTTTAT 660
QY 783 TATGAAGTCGCTCCAAATCTACAAACGAGTGTGATTCAGAGCTGTGCGGTATCA 842
DB 661 TATGAAGTCGCTCCAAATCTACAAACGAGTGTGATTCAGAGCTGTGCGGTATCA 720
QY 843 ATGCAATCTTCTGATTAATACCATCAACGAGAAAGTATTGTTTACACACAGCTAATGGC 902
DB 721 ATGCAATCTTCTGATTAATACCATCAACGAGAAAGTATTGTTTACACACAGCTAATGGC 780
QY 903 TACACCAATTAACCTAACCGTACCTACTCAAAAATAATACCAAAAAGGCTAGACCT 962
DB 781 TACACCAATTAACCTAACCGTACCTACTCAAAAATAATACCAAAAAGGCTAGACCT 840
QY 963 CTGCTACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 1022
DB 841 CTGCTACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 900
QY 1023 CTAGAAAAGCAATGATGCGCTCATTCG 1051
DB 901 CTAGAAAAGCAATGATGCGCTCATTCG 929
```

RESULT 4  
AAQ85037

ID AAQ85037 standard; DNA; 944 BP.

XX AC AAQ85037;

XX AC AAQ85037;

DT 25-MAR-2003 (updated)

DT 15-AUG-1995 (first entry)

```
XX DE DNA-ase-B gene for fusion to phage lambda promoter.
KW KW DNA-ase-B; phage lambda promoter; diagnostic; vaccine;
XX KW cystic fibrosis therapy; ds.
OS Streptococcus pyogenes (ATCC 14289).
XX Key Location/Qualifiers
FH CDS 1..942
FT /*tag= a
FT /product= DNA-ase-B
FT sig_peptide 1..135
FT /*tag= b
FT mat_peptide 136..939
FT /*tag= c
FT primer_bind complement (1..39)
FT /*tag= d
FT primer_bind 903..944
FT /*tag= e
FT misc_difference 763..765
FT /*tag= f
FT misc_difference 769..771
FT /*tag= g
FT /codon= seq:AGA, aa:Thr
XX MO9500650-A1.
XX 05-JAN-1995.
XX 18-MAY-1994; 94WO-US05626.
XX 23-JUN-1993; 93US-0082845.
XX (BECI ) BECKMAN INSTR INC.
XX Adams CW, Belei CM, Pang PPY;
XX WPI; 1995-052087/07.
XX P-PSDB; AAR/0702.
XX New DNA encoding Streptococcus pyogenes DNase B - for diagnosing
PT S. pyogenes infection, also new promoter for expressing other
PT proteins
XX Disclosure; Fig 5; 97pp; English.
XX The sequence shows a construct used in fusion of a phage lambda
CC promoter to DNA encoding a Streptococcus pyogenes DNA-ase-B.
CC Binding sites for polymerase chain reaction DNA primers (given
CC in AAQ85038 and AAQ85039) are shown. The product is useful in
CC production of diagnostic agents, vaccines or therapy of cystic
CC fibrosis.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 944 BP; 321 A; 191 C; 193 G; 239 T; 0 other;
Query Match 84.9%; Score 919.4; DB 16; Length 944;
Best Local Similarity 99.4%; Pred. No. 1.2e-228;
Matches 923; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAATGTCGGCTAGTAAA 182
DB 1 ATGGATCCGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAATGTCGGCTAGTAAA 60
QY 183 TTTTCAATGTAGTCTTGTATCAGCCCAATGGCTGTAAACAGTCACTTGAAT 242
DB 61 TTTTCAATGTAGTCTTGTATCAGCCCAATGGCTGTAAACAGTCACTTGAAT 120
QY 243 ACTGCATGCGCAGCAACACACAGCTCTCAATGATGTTGTTCTAAATGATCGCGCAAGC 302
DB 121 ACTGCATGCGCAGCAACACACAGCTCTCAATGATGTTGTTCTAAATGATCGCGCAAGC 180
```

```
QY 303 AAGTACCTAAACGAGCATTAGCTTGGACATTCATAGCAGCTCCTAACTATTACAAAAC 362
Db 181 AAGTACCTAAACGAGCATTAGCTTGGACATTCATAGCAGCTCCTAACTATTACAAAAC 240
QY 363 TTAGGTACTAGTACAGATTACTCCAGCAGCTTTCTTAAAGCAGGAGATATCTCTATAGC 422
Db 241 TTAGGTACTAGTACAGATTACTCCAGCAGCTTTCTTAAAGCAGGAGATATCTCTATAGC 300
QY 423 AAATTAGATGATTAGGAAGCAGCTACTCTAGAGGTACATTGACTTATGCCAATGTT 482
Db 301 AAATTAGATGATTAGGAAGCAGCTACTCTAGAGGTACATTGACTTATGCCAATGTT 360
QY 483 GAAGGTAGTACGGTGTAGCAATCTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 542
Db 361 GAAGGTAGTACGGTGTAGCAATCTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 420
QY 543 GGAACCTTAATCATGTCAAATATAAAATGAATGGTTAAATGCTCTATCTTATGTCGGA 602
Db 421 GGAACCTTAATCATGTCAAATATAAAATGAATGGTTAAATGCTCTATCTTATGTCGGA 480
QY 603 GATTTCGGAATAGAGTCACTCTCATTCGAGATAGTCTCGGTGGAGATGCACCTCAGAGTC 662
Db 481 GATTTCGGAATAGAGTCACTCTCATTCGAGATAGTCTCGGTGGAGATGCACCTCAGAGTC 540
QY 663 AATGCCGTTACAGGAACACGTCACCAAAATGTAGGAGTGTGACCAAAAAGGCGGCATG 722
Db 541 AATGCCGTTACAGGAACACGTCACCAAAATGTAGGAGTGTGACCAAAAAGGCGGCATG 600
QY 723 CGCTATACCGAACAAGAGCTCAGAAATGGTTAGCAACCAATCGTGATGGCTATCTTTAT 782
Db 601 CGCTATACCGAACAAGAGCTCAGAAATGGTTAGCAACCAATCGTGATGGCTATCTTTAT 660
QY 783 TATGAAGTCGCTCAATCTACAACGACAGAGTTGATTCAGAGCTGTCGTTGATCA 842
Db 661 TATGAAGTCGCTCAATCTACAACGACAGAGTTGATTCAGAGCTGTCGTTGATCA 720
QY 843 ATGCAATCTCTGATATACATCAACGAGAAAGTATTAGTTTACACACAGCTTAATGGC 902
Db 721 ATGCAATCTCTGATATACATCAACGAGAAAGTATTAGTTTACACACAGCTTAATGGC 780
QY 903 TACACCAATTAACCTACCAATAGCTTACACCTACTCAAAAATTAACCAAAAGGCTAGACCT 962
Db 781 TACACCAATTAACCTACCAATAGCTTACACCTACTCAAAAATTAACCAAAAGGCTAGACCT 840
QY 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 1022
Db 841 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 900
QY 1023 CTAGAAAAGCAATGATTGCCGTCATTGC 1051
Db 901 CTAGAAAAGCAATGATTGCCGTCATTGC 929

RESULT 5
ABN70648
ID ABN70648 standard; DNA; 813 BP.
XX
AC ABN70648;
XX
XX
DT 01-JUL-2002 (first entry)
XX
DE Streptococcus polynucleotide SEQ ID NO 9209.
XX
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX
OS Streptococcus pyogenes.
XX
XX
PN W0200234771-A2.
XX
PD 02-MAY-2002.
```

XX PF 29-OCT-2001; 2001WO-GB04789.  
XX PR 27-OCT-2000; 2000GB-0026333.  
XX PR 24-NOV-2000; 2000GB-0028727.  
XX PR 07-MAR-2001; 2001GB-0005640.  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX Telford J, Masignani V, Margarit Ros VI, Grandi G, Fraser C;  
PI Tettelin H;  
XX  
DR WPI; 2002-352536/38.  
DR P-PSDB; ABP30017.  
XX  
PT New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -  
XX  
PS Claim 7; Page 4044-4045; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
SQ Sequence 813 BP; 271 A; 165 C; 168 G; 209 T; 0 other;

Query Match 73.7%; Score 798.6; DB 24; Length 813;  
Best Local Similarity 98.9%; Pred. No. 2.6e-197;  
Matches 804; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 129 ATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAAAAATGTCGGCTAGTAAATTTTCA 188  
Db 1 ATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAAAAATGTCGGCTAGTAAATTTTCA 60  
QY 189 ATGCTAGCTCTTGTATCAGCCACAATGGCTGTAAACACAGTCACACTTGAATACTATGCA 248  
Db 61 ATGCTAGCTCTTGTATCAGCCACAATGGCTGTAAACACAGTCACACTTGAATACTATGCA 120  
QY 249 CTGGCAGCAAAACACACAGGTCCTCAAAATGATGTTGTTCTAAATGATGCGCGCAAGTAC 308  
Db 121 CTGGCAGCAAAACACACAGGTCCTCAAAATGATGTTGTTCTAAATGATGCGCGCAAGTAC 180  
QY 309 CTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGCTCCTTAATATTACAAACCTTAGGT 368  
Db 181 CTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGCTCCTTAATATTACAAACCTTAGGT 240  
QY 369 ACTAGTCAGATTACTCCAGCAGCTCTTTCTTAAAGCAGGAGATATCTCTATAGCAATTA 428  
Db 241 ACTAGTCAGATTACTCCAGCAGCTCTTTCTTAAAGCAGGAGATATCTCTATAGCAATTA 300  
QY 429 GATGAGTTAGGAAGGACGGCTACTGCTAGAGGTACATTGACTTATGCAATGTTGAAGGT 488  
Db 301 GATGAGTTAGGAAGGACGGCTACTGCTAGAGGTACATTGACTTATGCAATGTTGAAGGT 360  
QY 489 AGCTACGGTGTAGACAATCTTTCGGTAAAAATCAAAACCCCGCAGGATGAGCTGGAAC 548  
Db 361 AGCTACGGTGTAGACAATCTTTCGGTAAAAATCAAAACCCCGCAGGCTGAGCTGGAAC 420

Qy	549	CCTAATCATGTCAAATATATAAAATGCAATGGTTAAATGGTCTATCTTATGTCGGAGATTTC	608
Db	421	CCTAATCATGTCAAATATATAAAATGCAATGGTTAAATGGTCTATCTTATGTCGGAGATTTC	480
Qy	609	TGGAATAGAGTTCATCTCAITTCGAGATAGTCTTCGGTGGAGATGCACCTCAGAGTCAATGCC	668
Db	481	TGGAATAGAGTTCATCTCAITTCGAGATAGTCTTCGGTGGAGATGCACCTCAGAGTCAATGCC	540
Qy	669	GTTCACAGGAACGCTACCCCAAATGTAGAGGTGGTGCACCAAAAGCGCGCATGCCTAT	728
Db	541	GTTCACAGGACAGTACCCAAATGTAGAGGTGGTGCACCAAAAGCGCGCATGCCTAT	600
Qy	729	ACCGAAACAAAGACTCAAGATGGTTAGAACAAATCGTGATGGCTATCTTTATTATGAA	788
Db	601	ACCGAAACAAAGACTCAAGATGGTTAGAACAAATCGTGATGGCTATCTTTATTATGAA	660
Qy	789	GTGCTGCCAATCTCAACGCGAGAGGTTGATTCGAAGAGCTGCTGTTGATCAATGCAA	848
Db	661	GCTGCTCCANCTTATTAACGCGAGAGGTTGATTCGAAGAGCTGCTGTTGATCAATGCAA	720
Qy	849	TCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTAATGGCTACACC	908
Db	721	TCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTAATGGCTACACC	780
Qy	909	ATTAACTACCATTAACGGTACACCTACTCAAAA	941
Db	781	ATTAACTACCATTAACGGTACACCTACTCAGAA	813

RESULT 6	
AAQ85036	
ID	AAQ85036 standard; DNA; 400 BP.
XX	
AC	AAQ85036;
XX	
DT	25-MAR-2003 (updated)
DT	14-AUG-1995 (first entry)
XX	
DE	DNA-ase-B2 gene fragment.
XX	
KW	DNA-ase-B2 N-terminal fragment clone; diagnostic; ds.
XX	
OS	Streptococcus pyogenes.
XX	
PFH	Location/Qualifiers
CDS	129..400
FT	/tag= a
FT	/product= DNA-ase-B N-terminal fragment
XX	

```

CC infection.
CC (Updated on 25-MAR-2003 to correct PN field..)
XX
SQ Sequence 400 BP; 135 A; 77 C; 65 G; 123 T; 0 other;

Query Match      36.9%; Score 400; DB 16; Length
Best Local Similarity 100.0%; Pred. No. 5.9e-94;
Matches 400; Conservative 0; Mismatches 0; Indel:

QY      1  GACAAACGCCCTCTTTTTTCTCCCTACTACTCTCTCTTTAAATTTTCATATA
Db      1  GACAAACGCCCTCTTTTTTCTCCCTACTACTCTCTCTTTAAATTTTCATATA
QY     61  TATTGATAAACTAGTTTAAGTAAGCGGTATATCTATGGTTAGTTCAGCAACA
Db     61  TATTGATAAACTAGTTTAAGTAAGCGGTATATCTATGGTTAGTTCAGCAACA
QY    121  ACAAGCATATGANTCTACTTGGATCAAGACGGGTTTTTTCTAAAAAAA
Db    121  ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAAA
QY    181  AATTTTCAATGGTAGTCTTTGTATCAGCCACAATGGCTGTAACAACA
Db    181  AATTTTCAATGGTAGTCTTTGTATCAGCCACAATGGCTGTAACAACA
QY    241  ATACTGCACCTGGCAGCAAAAACAGGGTCTCAATGATGTTGTTCTTA
Db    241  ATACTGCACCTGGCAGCAAAAACAGGGTCTCAATGATGTTGTTCTTA
QY    301  GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCGAATGACAGTCTCTT
Db    301  GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCGAATGACAGTCTCTT
QY    361  CTTTAGGTACTAGTCAGATTAATCTCCAGACCTCTTTTCCTAA 400
Db    361  CTTTAGGTACTAGTCAGATTAATCTCCAGACCTCTTTTCCTAA 400

RESULT 7
AAT12773
ID AAT12773 standard; DNA; 400 BP.
AC AAT12773;
XX
XX
XX 25-JUN-1996 (first entry)
XX
XX S. pyogenes DNaseB DNA clone partial sequence.
XX DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;
KW diagnosis; Escherichia coli; ds.

```

```

OS Streptococcus pyogenes strain ATCC 14289.
XX
FH Key Location/Qualifiers
FT CDS 129..400
FT FT /*tag= a
FT FT sig_peptide 129..257
FT FT /*tag= b
FT FT mat_peptide 128..397
FT FT /*tag= c
XX
XX WO9606174-A1.
XX
XX
XX 29-FEB-1996.
XX
XX 18-AUG-1994; 94WO-US09450.
XX
XX 18-AUG-1994; 94WO-US09450.
XX
XX (BECI ) BECKMAN INSTR INC.
XX
XX Adams CW, Belei MC, Pang PPY;
XX

```





PD 29-FEB-1996.  
 XX 18-AUG-1994; 94WO-US09450.  
 XX 18-AUG-1994; 94WO-US09450.  
 XX (BECI ) BECKMAN INSTR INC.  
 XX Adams CW, Belei MC, Pang PPY;  
 XX WPI; 1996-151377/15.  
 XX New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
 PT Prodn. of the enzyme in other bacteria, useful in immunoassays or  
 PT for treating cystic fibrosis  
 XX Example 12; Page 54; 115pp; English.  
 PS A PCR primer pair (AAT12778-79) was used for the amplification of  
 CC Streptococcus pyogenes ATCC 14289 DNase B gene (see AAT12774).  
 CC Use of these primers gives a product that is processed  
 CC identically to native DNase B when inserted into vector plasmid  
 CC del-33 and expressed in Escherichia coli transformants. The  
 CC N-terminal sequence of the recombinant DNase B is identical to  
 CC that of the native mature enzyme (AAR8824). Highly purified  
 CC DNase B can be obtd. in large quantities.  
 XX Sequence 182 BP; 40 A; 49 C; 42 G; 51 T; 0 other;  
 SQ  
 Query Match 9.8%; Score 106.6; DB 17; Length 182;  
 Best Local Similarity 77.3%; Pred. No. 6.5e-18;  
 Matches 143; Conservative 0; Mismatches 39; Indels 3; Gaps 1;  
 QY 118 AGGACAGCATATGATCTACTTGGATCAAGACGGGTTTTCTTAAATAATCGGCTAG 177  
 DB 1 AGGCAATGGATCGGAACCTGCTGGGTCCCGTGGTTTCTCCAAAATGCCGCTGG 60  
 QY 178 TAAATTTTCAATGGTAGCTCTGTATCAGCCACATGGCTGTAAACAACAGTCACATTG 237  
 DB 61 TTAATTTCCATGGTGTCTGTGTTTCCGCTACCATGGCTGTACCACCGTTACCGTGG 120  
 QY 238 AAATATCTGCTGGCAGCAGCAAAACACAGGTCTCAAATGATGTTCTTAAATGATGGCG 297  
 DB 121 AAACACCGCTGTGGC---TCAGACACAGGTCTCAAATGATGTTCTTAAATGATGGCG 177  
 QY 298 CAAGC 302  
 DB 178 CAAGC 182  
 RESULT 10  
 AAK99137  
 ID AAK99137 standard; DNA; 966 BP.  
 AC AAK99137;  
 XX 12-JUN-2002 (first entry)  
 DT 966nt DNA sequence of the invention.  
 DE Recombinant streptodornase; mutated Streptococcus equisimilis;  
 KW mass production; ds.  
 KW Unidentified.  
 OS KR99041925-A.  
 PN 15-JUN-1999.  
 PD 25-NOV-1997; 97KR-0062603.  
 PF 25-NOV-1997; 97KR-0062603.  
 PR 25-NOV-1997; 97KR-0062603.  
 XX

PA (LEEH/) LEE H H.  
 XX Bae S; Kim IC, Lee HH, Sohn HJ;  
 PI WPI; 2000-408398/35.  
 DR Recombinant streptodornase and process for its mass production from  
 XX mutated Streptococcus equisimilis -  
 XX Disclosure; Page 3; 11pp; Korean.  
 PS The invention relates to a recombinant streptodornase and a process for  
 CC its mass production from mutated Streptococcus equisimilis. This  
 CC polynucleotide sequence represents a 966nt DNA of the invention.  
 XX Sequence 966 BP; 363 A; 157 C; 184 G; 262 T; 0 other;  
 SQ  
 Query Match 8.2%; Score 88.4; DB 21; Length 966;  
 Best Local Similarity 55.4%; Pred. No. 6.1e-13;  
 Matches 195; Conservative 0; Mismatches 151; Indels 6; Gaps 1;  
 QY 579 TTAATGGTCTATCTTATGTCGAGATTTCTGGAATAGAAAGTCATCTCATTCAGATAGT 638  
 DB 436 TTAAGGTAAATAATACACAGTCATCTTATTTGTAGCTAGCCATCTCTTTCCGATAGC 495  
 QY 639 CTCGGTGGAGATGCACTCAGAGTCATGCGTTTACAGGAACACGTCACCCAAATGTAGGA 698  
 DB 496 CTTCGAGGAAAGTCCATACGGAATAATGCTATTCTGGAATCAATGCAAAATGTGGG 555  
 QY 699 GGTCTGTGCAAAAGGCGGATGCGCTATACCGAACAAAGAGCTCAAGAATGGTTAGAA 758  
 DB 556 ACAGT-----AAAGTGGGATGCAATATATTGAAAAAAGTTTAAAGCCATATCACT 609  
 QY 759 GCAATTCGTGATGGCTATCTTTATTATGAGTCGCTCCAATCTACAACGACAGAGTTG 818  
 DB 610 AAGATTCCTGATGTTTATGTTATCTACAGTCCATCCCTGAATATCAGGGGGCTGAGTTA 669  
 QY 819 ATTCCAGAGTGTGCTGGTATCAATGCAATCTTCTGATATATACCATCAACGAGAAAGTA 878  
 DB 670 TTAGCAAGATCAGTTTATGATATCTGCTTTATCTTCGATGGAGTAATTATGAACCTGTT 729  
 QY 879 TTAGTTTCAACACAGTAAATGGCTACACCAATTAACCTACCATACCGGTACAC 930  
 DB 730 CGTGTTCACACCGCTGATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 781  
 RESULT 11  
 ABN91441  
 ID ABN91441 standard; DNA; 2409 BP.  
 XX AC ABN91441;  
 XX 24-JUL-2002 (first entry)  
 DT Staphylococcus epidermidis ORF nucleic acid sequence SEQ ID NO:904.  
 DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;  
 KW antibacterial; gene therapy; gene; ds.  
 KW Staphylococcus epidermidis.  
 OS US6380370-B1.  
 PN 30-APR-2002.  
 PD 13-AUG-1998; 98US-0134001.  
 PF 14-AUG-1997; 97US-055779P.  
 PR 08-NOV-1997; 97US-064964P.  
 XX (GENO-) GENOME THERAPEUTICS CORP.  
 PA Doucette-Stamm LA, Bush D;  
 PI





[illegible]

RESULT 15	
ABN70213/c	
ID	ABN70213 standard; DNA; 126 BP.
XX	XX
AC	ABN70213;
XX	XX
DT	01-JUL-2002 (first entry)
XX	XX
DE	Streptococcus polynucleotide SEQ ID NO 8339.
XX	XX
KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW	group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
KW	antiinflammatory; infection; vaccine; meningitis; gene therapy; ds.
XX	XX
OS	Streptococcus pyogenes.
XX	XX
PN	W0200234771-A2.
XX	XX
PD	02-MAY-2002.
XX	XX
PF	29-OCT-2001; 2001WO-GB04789.
XX	XX
PR	27-OCT-2000; 2000GB-0026333.
PR	24-NOV-2000; 2000GB-0028727.
PR	07-MAR-2001; 2001GB-0005640.
XX	XX
PA	(CHIR-) CHIRON SPA.
PA	(GENO-) INST GENOMIC RES.
XX	XX
PI	Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
PI	Tettelin H;
XX	XX
WPI	WPI; 2002-352536/38.
DR	P-PSDB; ABP29582.
XX	XX
PT	New Streptococcus protein for the treatment or prevention of infection
PT	or disease caused by Streptococcus bacteria, such as meningitis, and
PT	for detecting a compound that binds to the protein .
XX	XX
PS	Claim 7; Page 3948; 4525pp; English.
XX	XX
CC	The invention relates to a protein (ABP25413-ABP30895) from group B
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC	the specification. The proteins have antibacterial and antiinflammatory

CC	activity. (I), nucleic acids encoding (I), ABB6044-ABN71526 and
CC	antibodies that bind (I) are used in the manufacture of medicaments for
CC	the treatment or prevention of infection or disease caused by
CC	<i>Streptococcus</i> bacteria, particularly <i>S. agalactiae</i> and <i>S. pyogenes</i> .
CC	Nucleic acids encoding (I) are used to detect <i>Streptococcus</i> in a
CC	biological sample. (I) is used to determine whether a compound binds to
CC	(I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC	used as a vaccine or diagnostic composition. The disease caused by
CC	<i>Streptococcus</i> that is prevented or treated may be meningitis. Nucleic
CC	acid encoding (I) may be used to recombinantly produce (I) and may be
CC	used in gene therapy. Antibodies to (I) are used for affinity
CC	chromatography, immunoassays, and distinguishing/identifying
CC	<i>Streptococcus</i> proteins.
XX	
SQ	Sequence 126 BP; 39 A; 12 C; 33 G; 42 T; 0 other;
	Query Match            4.2%; Score 45; DB 24; Length 126;
	Best Local Similarity 100.0%; Pred. No. 0.055;
	Matches         45; Conservative         0; Mismatches         0; Indels         0; Gaps         0;
Qy	1 GACAAAGCCTCTTTTTCCTACTATCTCCTTAATTTCAT 45
Dd	45 GACAAAGCCTCTTTTTCCTACTATCTCCTTAATTTCAT 1

Search completed: January 5, 2004, 14:01:44  
Job time : 354 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:53:48 ; Search time 92 Seconds  
(without alignments)  
5195.847 Million cell updates/sec

Title: US-08-482-785-7  
Perfect score: 1083  
Sequence: 1 GACAAAGCCTCTCTTTTCT.....TGTCAAAAGCAAAAGC 1083

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/2/ina/5A-COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B-COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A-COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B-COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS-COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1083	100.0	1083	4	US-08-393-889-7
2	1083	100.0	1083	5	PCT-US94-09450-7
3	922.6	85.2	940	4	US-08-393-889-11
4	922.6	85.2	940	5	PCT-US94-09450-11
5	853.8	78.8	937	4	US-08-393-889-14
6	853.8	78.8	937	5	PCT-US94-09450-14
7	200	18.5	200	4	US-08-393-889-10
8	200	18.5	200	5	PCT-US94-09450-10
9	106.6	9.8	182	4	US-08-393-889-12
10	106.6	9.8	182	5	PCT-US94-09450-12
11	47.2	4.4	2409	4	US-09-134-001C-904
12	45	4.2	7218	1	US-08-232-463-14
13	37.4	3.5	5703	4	US-09-280-590A-36
14	37.2	3.4	1844	3	US-08-134-557D-1
15	36.4	3.4	5333	4	US-09-623-062-1
16	36.2	3.3	588	4	US-08-887-534A-56
17	36.2	3.3	588	4	US-09-527-431-56
18	35.2	3.3	12720	1	US-08-403-866-11
19	35	3.2	588	4	US-08-887-534A-58
20	35	3.2	588	4	US-09-527-431-58
21	34.8	3.2	5238	2	US-08-521-053-11
22	34.8	3.2	5688	4	US-09-546-934-3
23	34.6	3.2	495	4	US-09-220-132-186
24	34.6	3.2	732	4	US-09-328-352-2414
25	34.6	3.2	3001	4	US-09-539-333D-158
26	34.6	3.2	3001	4	US-09-539-333D-160
27	34.6	3.2	168575	4	US-09-426-290-1

C 28 34.6 3.2 1664976 4 US-08-916-421B-1 Sequence 1, Appli  
29 34.4 3.2 665 2 US-08-883-795A-36 Sequence 36, Appl  
30 34.4 3.2 4215 4 US-09-620-312B-295 Sequence 295, App  
C 31 34.2 3.2 940 4 PCT-US94-09450-11 Sequence 11, Appl  
32 34.2 3.2 940 5 PCT-US94-09450-11 Sequence 11, Appl  
C 33 34.2 3.2 1083 4 US-08-393-889-7 Sequence 7, Appli  
34 34.2 3.2 1083 5 PCT-US94-09450-7 Sequence 7, Appli  
C 35 34.2 3.2 2235 4 US-09-569-804-20 Sequence 20, Appl  
36 34.2 3.2 5852 1 US-07-867-106-2 Sequence 2, Appli  
C 37 34.2 3.2 15788 4 US-09-920-759-13 Sequence 13, Appl  
C 38 34 3.1 1990 4 US-08-961-527-232 Sequence 871, App  
C 39 34 3.1 4016 4 US-09-173-053-3 Sequence 3, Appli  
C 40 34 3.1 602 1 US-08-764-100-8 Sequence 8, Appli  
C 41 33.4 3.1 642 1 US-08-764-100-13 Sequence 13, Appl  
42 33.4 3.1 643 1 US-08-764-100-7 Sequence 13, Appli  
C 43 33.4 3.1 1257 4 US-09-134-001C-1214 Sequence 1214, Ap  
44 33.4 3.1 1395 1 US-07-991-867B-25 Sequence 25, Appl  
C 45 33.4 3.1 1395 1 US-07-991-867B-25 Sequence 25, Appl

#### ALIGNMENTS

RESULT 1  
US-08-393-889-7  
; Sequence 7, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1083 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: 129..944  
US-08-393-889-7

```
Query Match      100.0%; Score 1083; DB 4; Length 1083;
Best Local Similarity 100.0%; Pred. No. 7.8e-288;
Matches 1083; Conservative 0; Mismatches 0; Indels 0;
```

QY		1	GACAAAGCGCTTCCTTTTCTCCTTACTATCTCCCTTAATTTTCATATTTTTTAAAAAAC	60
Db		1	GACAAAGCGCTTCCTTTTCTCCTTACTATCTCCCTTAATTTTCATATTTTTTAAAAAAC	60
QY		61	TATTGATAAATAGTTAAGTAAAGCGTATACCTATGTTAGTTAGCGAAATTAGAAGAAGG	120
Db		61	TATTGATAAATAGTTAAGTAAAGCGTATACCTATGTTAGTTAGCGAAATTAGAAGAAGG	120
QY		121	ACAAGCATATGAATCTACTTGGAATCAAGACGGGTGTTTTCTPAAAAAATGTCGGCTAGTAA	180
Db		121	ACAAGCATATGAATCTACTTGGAATCAAGACGGGTGTTTTCTPAAAAAATGTCGGCTAGTAA	180
QY		181	AATTTTCAAATGGTAGCTCTTGTTATCAGGCCAAATGGCTGTAAACAACAGTCACTTGA	240
Db		181	AATTTTCAAATGGTAGCTCTTGTTATCAGGCCAAATGGCTGTAAACAACAGTCACTTGA	240
QY		241	ATACTGCACCTGGCACGACAAAACAAGCTCTCAAAATGATGTTGTTCTAAATGATCGCGCAA	300
Db		241	ATACTGCACCTGGCACGACAAAACAAGCTCTCAAAATGATGTTGTTCTAAATGATCGCGCAA	300
QY		301	GCAAGTACCCTAAAGCAAGCATTAGCTTGGACATTCATATGACAGTCTCTAATTCACAAA	360
Db		301	GCAAGTACCCTAAAGCAAGCATTAGCTTGGACATTCATATGACAGTCTCTAATTCACAAA	360
QY		361	CTTTAGGTACTAGTCAGATTACTCCAGCACTCTTCTCTAAAGCAGGAGATATCTCTATA	420
Db		361	CTTTAGGTACTAGTCAGATTACTCCAGCACTCTTCTCTAAAGCAGGAGATATCTCTATA	420
QY		421	GCAAAATTAGATAGTTAGGAAGACGGCTACTGCTAGAGGTACATTCGACTTATGCCAATG	480
Db		421	GCAAAATTAGATAGTTAGGAAGACGGCTACTGCTAGAGGTACATTCGACTTATGCCAATG	480
QY		481	TTGAAGGTACTACGGGTGTAGACAAATCTTTTCGTTAAATAATCAAAACCCCGCAGGATGGA	540
Db		481	TTGAAGGTACTACGGGTGTAGACAAATCTTTTCGTTAAATAATCAAAACCCCGCAGGATGGA	540
QY		541	CTGGAAACCCCTAATCTATGTCATAATATAAAATGAAATGGTTTAAATGGTCTATCTTATGTCG	600
Db		541	CTGGAAACCCCTAATCTATGTCATAATATAAAATGAAATGGTTTAAATGGTCTATCTTATGTCG	600
QY		601	GAGATTTCTGGAATAGAAAGTCATCTCATTCCAGATAGTCTCGTGGAGATGCACTCAGAG	660
Db		601	GAGATTTCTGGAATAGAAAGTCATCTCATTCCAGATAGTCTCGTGGAGATGCACTCAGAG	660
QY		661	TCAATCCGGTTACAGGAACAGTACCCAAAATGTAGGAGTCTGTGACCAAAAAAGGCGCCA	720
Db		661	TCAATCCGGTTACAGGAACAGTACCCAAAATGTAGGAGTCTGTGACCAAAAAAGGCGCCA	720
QY		721	TGCGCTATACCGAAACAAAGACTCAAGAAATGGTTTAGAACAAATCGTGATGGCTATCTTT	780
Db		721	TGCGCTATACCGAAACAAAGACTCAAGAAATGGTTTAGAACAAATCGTGATGGCTATCTTT	780
QY		781	ATTATGAAGTCGCTCCAAATCTCAACGACAGCAGTTGATTTCAAGAGCTGTCGTGGTAT	840
Db		781	ATTATGAAGTCGCTCCAAATCTCAACGACAGCAGTTGATTTCAAGAGCTGTCGTGGTAT	840
QY		841	CMAATGAATCTTCTGATAATACCATCAACGAGAAAAGTATTAGTTTACAACACAGCTAAATG	900
Db		841	CMAATGAATCTTCTGATAATACCATCAACGAGAAAAGTATTAGTTTACAACACAGCTAAATG	900
QY		901	GCTACACCAATTACTACCTAAGCGGTACCTACTCAAAATATATACCAAAAGGCTAGAC	960
Db		901	GCTACACCAATTACTACCTAAGCGGTACCTACTCAAAATATATACCAAAAGGCTAGAC	960
QY		961	CTCTGCTCTACTAGGCCCTAGCTTTTTCATCAAAAAAGCAATGACTATATGAAATGAAAA	1020

Db	961	CTCTGCTC	TACTAGGCCTAGCTTTT	TATCATCAAAAAAGCAATGACTATAGAAAGTAAAAA	1020
Qy	1021	TACTAGAAAAAGCAATGATTGCCGTCATTTGCTTTT	TATGAAATTTGTGCAAAAAAGCAAAAAA	1080	
Db	1021	TACTAGAAAAAGCAATGATTGCCGTCATTTGCTTTT	TATGAAATTTGTGCAAAAAAGCAAAAAA	1080	
Qy	1081	AGC	1083		
Db	1081	AGC	1083		

## RESULT 2

```

PCT-US94-09450-7
; Sequence 7, Application PC/TUS9409450
; GENERAL INFORMATION:
; APPLICANT: Beckman Instruments, Inc.
; APPLICANT: 2500 Harbor Boulevard
; APPLICANT: Fullerton, California 92634
; TITLE OF INVENTION: Recombinant Dnase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beckman Instruments, Inc.
; STREET: 2500 Harbor Boulevard
; CITY: Fullerton
; STATE: California
; COUNTRY: USA
; ZIP: 92634
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09450
; FILING DATE: 18-AUG-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/082,845
; FILING DATE: 23-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: May, William H.
; REGISTRATION NUMBER: 26,769
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1083 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 129..944
PCT-US94-09450-7

```

	Query Match	100.0%	Score 1083;	DB 5;	Length 1083;
	Best Local Similarity	100.0%;	Pred. No. 7.8e-288;		
	Matches 1083;	Conservative	0;	Mismatches	0; Indels
					0; Gaps
QY	1	GACACGGCCTCTTTTTCTCCTTACTATCCTCTTAATTTTCATATTTTAAAAAAC	60		
Db	1	GACACGGCCTCTTTTTCTCCTTACTATCCTCTTAATTTTCATATTTTAAAAAAC	60		
QY	61	TATTGATAAACCTAGTTAAAGTAAAGCGTATATCATGTGGTTAGTAGCGAAATTAGAAAAGAGG	120		



Db 61 TATTGATAAAGTAAAGCGGTACTATGTTAGTTCGCGAAATAGAAAAGAGG 120  
QY 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAATGTCGCTAGTAA 180  
Db 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAATGTCGCTAGTAA 180  
QY 181 AATTTTCAATGGTGTCTGTATCAGCCACATGGCTGTAAACAGTCACACTTTGAAA 240  
Db 181 AATTTTCAATGGTGTCTGTATCAGCCACATGGCTGTAAACAGTCACACTTTGAAA 240  
QY 241 ATACTGCACCTGGCAGCAGCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCGCAA 300  
Db 241 ATACTGCACCTGGCAGCAGCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCGCAA 300  
QY 301 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGCTCTTAACCTATTACAAA 360  
Db 301 GCAAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGCTCTTAACCTATTACAAA 360  
QY 361 CTTTAGGTACTAGTCAAGTACTCCAGCACTCTTCTTAAGCAGGAGATATCTCTATA 420  
Db 361 CTTTAGGTACTAGTCAAGTACTCCAGCACTCTTCTTAAGCAGGAGATATCTCTATA 420  
QY 421 GCAAAATAGATGATTTAGGAAGACGCTACTCTGCTAGAGTACATTGACTTATGCCAATG 480  
Db 421 GCAAAATAGATGATTTAGGAAGACGCTACTCTGCTAGAGTACATTGACTTATGCCAATG 480  
QY 481 TTGAAGTACTAGTACGGTGTGTAGCAATCTTTTCGGTAAATCAAAACCCCGCAGATGGA 540  
Db 481 TTGAAGTACTAGTACGGTGTGTAGCAATCTTTTCGGTAAATCAAAACCCCGCAGATGGA 540  
QY 541 CTGGAACCTTAATCATGTCAATATAAATTAATGAATGGTAAATGCTCTATCTTATGTCG 600  
Db 541 CTGGAACCTTAATCATGTCAATATAAATTAATGAATGGTAAATGCTCTATCTTATGTCG 600  
QY 601 GAGATTTCTGGAATAGATGATCTCTCATTCGACATAGTCTCGTGAGAGTGCACCTCAGAG 660  
Db 601 GAGATTTCTGGAATAGATGATCTCTCATTCGACATAGTCTCGTGAGAGTGCACCTCAGAG 660  
QY 661 TCAATGCCGTGTACAGGAACAGCTACCAAAATGTAGGAGTCTGTGACCAAAAGCGCGCA 720  
Db 661 TCAATGCCGTGTACAGGAACAGCTACCAAAATGTAGGAGTCTGTGACCAAAAGCGCGCA 720  
QY 721 TGGCTTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Db 721 TGGCTTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
QY 781 ATTATGAAGTCTGCTCAATCTCAACGACAGAGTTGATTCAGAGCTGTGCTGTAT 840  
Db 781 ATTATGAAGTCTGCTCAATCTCAACGACAGAGTTGATTCAGAGCTGTGCTGTAT 840  
QY 841 CAATGCAATCTTCTGATATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 900  
Db 841 CAATGCAATCTTCTGATATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 900  
QY 901 GCTACACCATTAATACCTAACCGGTACACCTACTCAAAAATTAATACCAAAAGGCTAGAC 960  
Db 901 GCTACACCATTAATACCTAACCGGTACACCTACTCAAAAATTAATACCAAAAGGCTAGAC 960  
QY 961 CTCGTCTACTAGGCGCTAGCTTTTACATCAAAAAGCAATGACATATAGAAAAGTAAAAA 1020  
Db 961 CTCGTCTACTAGGCGCTAGCTTTTACATCAAAAAGCAATGACATATAGAAAAGTAAAAA 1020  
QY 1021 TACTAGAAAAGCAATGATGGCGTCAATGCTTTTATGAAATTTGTCAAAAGCAAAA 1080  
Db 1021 TACTAGAAAAGCAATGATGGCGTCAATGCTTTTATGAAATTTGTCAAAAGCAAAA 1080  
QY 1081 AGC 1083  
Db 1081 AGC 1083

RESULT 3

US-08-393-889-11

; Sequence 11, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 940 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; US-08-393-889-11

Query Match 85.2%; Score 922.6; DB 4; Length 940;  
Best Local Similarity 99.6%; Pred. No. 7.8e-244;  
Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 123 AAGCATATCAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAATGTCGCTAGTAAAA 182  
Db 1 ATGGATCCGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAATGTCGCTAGTAAAA 60  
QY 183 TTTTCAATGGTGTAGTCTCTGTATCAGCCACAAATGGCTGTAAACAGTCACACTTTGAAAAT 242  
Db 61 TTTTCAATGGTGTAGTCTCTGTATCAGCCACAAATGGCTGTAAACAGTCACACTTTGAAAAT 120  
QY 243 ACTGCACTGGCAGCAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGCAAGC 302  
Db 121 ACTGCACTGGCAGCAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGCAAGC 180  
QY 303 AAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCAATATTACAAACT 362  
Db 181 AAGTACCTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCAATATTACAAACT 240  
QY 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGGAGATTTCTCTATAGC 422  
Db 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGGAGATTTCTCTATAGC 300  
QY 423 AAATTAGATGAGTTAGGAAGGACGCGTACTGCTAGAGGTACATTGACTTATGCAATGTTT 482

```
Db 301 AAATTAGATGTTAGGAAGGACGCTACTCTAGAGGTACATTGACTTATGCGCAATGTT 360
Qy 483 GAAGTAGTACGGTGTAGCAATCTTTCCGTAAGAAATCAAAACCCGCGAGGTGACT 542
Db 361 GAAGTAGTACGGTGTAGCAATCTTTCCGTAAGAAATCAAAACCCGCGAGGTGACT 420
Qy 543 GGAACCCCTAATCATGTCAAAATATAAAATGAATGGTCTATCTTATGTCGGA 602
Db 421 GGAACCCCTAATCATGTCAAAATATAAAATGAATGGTCTATCTTATGTCGGA 480
Qy 603 GATTTCGGAATAGAGTCAATCTCATTCGATAGTCTCGGTGAGATGCACTCAGAGTC 662
Db 481 GATTTCGGAATAGAGTCAATCTCATTCGATAGTCTCGGTGAGATGCACTCAGAGTC 540
Qy 663 AATCCGCTTACAGGAACACGTPACCAAAATCTAGAGGTCTGTGACCAAAAGGCGGCATG 722
Db 541 AATCCGCTTACAGGAACACGTPACCAAAATCTAGAGGTCTGTGACCAAAAGGCGGCATG 600
Qy 723 CGCTATACCGAACAAGAGCTCAAGAATGGTTAGAAGCAAAATCGTGATGCTATCTTTAT 782
Db 601 CGCTATACCGAACAAGAGCTCAAGAATGGTTAGAAGCAAAATCGTGATGCTATCTTTAT 660
Qy 783 TATGAAGTCGCTCAATCTCAACGCGAGACGAGTTGATTCAGAGCTGTCGTGGTATCA 842
Db 661 TATGAAGTCGCTCAATCTCAACGCGAGACGAGTTGATTCAGAGCTGTCGTGGTATCA 720
Qy 843 ATGCAATCTCTGATATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTTAATGGC 902
Db 721 ATGCAATCTCTGATATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTTAATGGC 780
Qy 903 TACACCAATTAACCTACCAAGGTCACCTACTCAAAAATAATACCAAAAGGCTTAGACCT 962
Db 781 TACACCAATTAACCTACCAAGGTCACCTACTCAAAAATAATACCAAAAGGCTTAGACCT 840
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 1022
Db 841 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 900
Qy 1023 CTGAAAAAGCAATGATTGCGCTCATTCG 1051
Db 901 CTGAAAAAGCAATGATTGCGCTCATTCG 929
```

## RESULT 4

PCT-US94-09450-11

Sequence 11. Application PC/TUS9409450

GENERAL INFORMATION:

APPLICANT: Beckman Instruments, Inc.

APPLICANT: 2500 Harbor Boulevard

APPLICANT: Fullerton, California 92634

TITLE OF INVENTION: Recombinant DNase B Derived from

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESSES: 16

ADDRESSEE: Beckman Instruments, Inc.

STREET: 2500 Harbor Boulevard

CITY: Fullerton

STATE: California

COUNTRY: USA

ZIP: 92634

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09450

FILING DATE: 18-AUG-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/082,845

FILING DATE: 23-JUN-1993

```
ATTORNEY/AGENT INFORMATION:
NAME: May, William H.
REGISTRATION NUMBER: 26,769
REFERENCE/DOCKET NUMBER: 39D-1357 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 940 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
PCT-US94-09450-11
```

Query Match 85.2%; Score 922.6; DB 5; Length 940;

Best Local Similarity 99.6%; Pred. No. 7.8e-244;

Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAAAA 182
Db 1 ATGGATCCCGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAAAA 60
Qy 183 TTTTCAATGGTAGTCTTGTATCAGCCACAATGGCTGTAAACAACAGTCACACTTGAAT 242
Db 61 TTTTCAATGGTAGTCTTGTATCAGCCACAATGGCTGTAAACAACAGTCACACTTGAAT 120
Qy 243 ACTGCACTGGCAGCAAAACACAGGTCCTCAAAATGATGTTGTTCTAAATGATGCGCAAGC 302
Db 121 ACTGCACTGGCAGCAAAACACAGGTCCTCAAAATGATGTTGTTCTAAATGATGCGCAAGC 180
Qy 303 AAGTACTTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGTCCTAACTATTACAAACT 362
Db 181 AAGTACTTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGTCCTAACTATTACAAACT 240
Qy 363 TTAGGTACTAGTCAGATTAATCTCAGCAGCTCTTCTTCTTAAAGAGGAGATATTCTCTATAGC 422
Db 241 TTAGGTACTAGTCAGATTAATCTCAGCAGCTCTTCTTCTTAAAGAGGAGATATTCTCTATAGC 300
Qy 423 AAATTAGATGAGTTAGGAAGGACGCTACTGCTAGAGGTACATTGACTTTATGCAATGTT 482
Db 301 AAATTAGATGAGTTAGGAAGGACGCTACTGCTAGAGGTACATTGACTTTATGCAATGTT 360
Qy 483 GAAGGTAGCTACGGTGTAGACAAATCTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 542
Db 361 GAAGGTAGCTACGGTGTAGACAAATCTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 420
Qy 543 GGAACCCCTAATCATGTCAAAATATAAAATGAATGGTAAATGGTCTATCTTATGTCGGA 602
Db 421 GGAACCCCTAATCATGTCAAAATATAAAATGAATGGTAAATGGTCTATCTTATGTCGGA 480
Qy 603 GATTTCTGGAATAGAGTCAATCTCTATTGAGATAGTCTCGGTGGAGATGCACACTCAGAGTC 662
Db 481 GATTTCTGGAATAGAGTCAATCTCTATTGAGATAGTCTCGGTGGAGATGCACACTCAGAGTC 540
Qy 663 AATGCGGTTACAGGAACACGTCATCCCAAAATATAGGAGGTCTGTGACCAAAAGGCGGCATG 722
Db 541 AATGCGGTTACAGGAACACGTCATCCCAAAATATAGGAGGTCTGTGACCAAAAGGCGGCATG 600
Qy 723 CGCTATACCGAACAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGCTATCTTTAT 782
Db 601 CGCTATACCGAACAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGCTATCTTTAT 660
Qy 783 TATGAAGTCGCTCCAAATCTTACAAACGAGACGAGTTGATTCGAAGAGCTCTCGTGGTATCA 842
Db 661 TATGAAGTCGCTCCAAATCTTACAAACGAGACGAGTTGATTCGAAGAGCTCTCGTGGTATCA 720
Qy 843 ATGCAATCTTCTGATATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTTAATGGC 902
```

Db 721 ATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAAACACAGCTAATGGC 780  
Qy 903 TACACCAATTAACCTAACGAGTACACCTACTCAAAATAATAACCAAAAGGCTAGACCT 962  
Db 781 TACACCAATTAACCTAACGAGTACACCTACTCAAAATAATAACCAAAAGGCTAGACCT 840  
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAATAAGCAATGACTATAGAAAGTAAATA 1022  
Db 841 CTGCTCACTAGGCTAGCTTTTACATCAAAATAAGCAATGACTATAGAAAGTAAATA 900  
Qy 1023 CTAGAAAAGCAATGATTGCGCTCATGTC 1051  
Db 901 CTAGAAAAGCAATGATTGCGCTCATGTC 929

## RESULT 5

US-08-393-889-14  
; Sequence 14, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 937 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..819  
; US-08-393-889-14

Query Match 78.8%; Score 853.8; DB 4; Length 937;  
Best Local Similarity 95.7%; Pred. No. 66-225;  
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAAAAATGTCGGCTAGTAAA 182  
Db 1 ATGGATCCGAACCTGCTGGGTTCCCGTGGTGTTCCTCCAAAAAATGCGGCTGGTTAAA 60  
Qy 183 TTTTCAATGTGTAGTCTCTTGTATCAGCCACACATGGCTGTAAACAACAGTCACACTTGAATAAT 242  
Db 61 TTCTCCATGGTGTCTCTGGTTTCGGTACCATGGCTGTACACCGTTACCTCGGAAC 120  
Qy 243 ACTGCACTGGCAGCAAAACACACAGGTCTCAATATGATGTTTCTTAAATGATGGCGCAAGC 302  
Db 121 ACCGCTCTGGC---TCAGACACAGGTCTCAATATGATGTTTCTTAAATGATGGCGCAAGC 177  
Qy 303 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGTCTCTACTATTACAAACT 362  
Db 178 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGTCTCTACTATTACAAACT 237  
Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTCTATAGC 422  
Db 238 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTCTATAGC 297  
Qy 423 AAATTAGATGATTAGGAAGGACCGCTACTGCTAGAGGTACATTGACTTTATGCCAATGTT 482  
Db 298 AAATTAGATGATTAGGAAGGACCGCTACTGCTAGAGGTACATTGACTTTATGCCAATGTT 357  
Qy 483 GAAGGTAGCTACGGTGTAGCAATCTTTCCGTTAAAAATCAAAACCCCGCAGGATGACT 542  
Db 358 GAAGGTAGCTACGGTGTAGCAATCTTTCCGTTAAAAATCAAAACCCCGCAGGATGACT 417  
Qy 543 GGAACCCCTAATCATGTCAATATAAAATGAATGTTAAATGTTCTTATGTCGGA 602  
Db 418 GGAACCCCTAATCATGTCAATATAAAATGAATGTTAAATGTTCTTATGTCGGA 477  
Qy 603 GATTTCCTGGAATAGAAAGTCAATCTCATTCAGATAGTCTCGTGGAGATGCACTCAGAGTC 662  
Db 478 GATTTCCTGGAATAGAAAGTCAATCTCATTCAGATAGTCTCGTGGAGATGCACTCAGAGTC 537  
Qy 663 AATGCCGTTACAGAAACACGTACCCAAAATGTAGAGGTCTGTACCAAAAAGCGCGCATG 722  
Db 538 AATGCCGTTACAGAAACACGTACCCAAAATGTAGAGGTCTGTACCAAAAAGCGCGCATG 597  
Qy 723 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAACAAATCGTGATGGCTATCTTTAT 782  
Db 598 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAACAAATCGTGATGGCTATCTTTAT 657  
Qy 783 TATGAAGTCTGCTCCAATCTACACGACGAGCTTGATTCGAAGAGCTGTCGTGGTATCA 842  
Db 658 TATGAAGTCTGCTCCAATCTACACGACGAGCTTGATTCGAAGAGCTGTCGTGGTATCA 717  
Qy 843 ATGCAATCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGGC 902  
Db 718 ATGCAATCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGGC 777  
Qy 903 TACACCAATTAACCTACCAACGCTACCTACTCAAAATAATAACCAAAAGGCTAGACCT 962  
Db 778 TACACCAATTAACCTACCAACGCTACCTACTCAAAATAATAACCAAAAGGCTAGACCT 837  
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAATAAGCAATGACTATAGAAAGTAAATA 1022  
Db 838 CTGCTCACTAGGCTAGCTTTTACATCAAAATAAGCAATGACTATAGAAAGTAAATA 897  
Qy 1023 CTAGAAAAGCAATGATTGCGCTCATGTC 1051  
Db 898 CTAGAAAAGCAATGATTGCGCTCATGTC 926

## RESULT 6

PCT-US94-09450-14  
; Sequence 14, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 937 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..819  
PCT-US94-09450-14

Query Match 78.8%; Score 853.8; DB 5; Length 937;  
Best Local Similarity 95.7%; Pred. No. 6e-225;  
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

QY 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTCTTAAATAATGTCGGCTAGTAA 182  
Db 1 ATGGATCCGAACCTGCTGGGTTCCCGTCTGTTTCTCCAAAAAATGCGGTCGTGTTAA 60

QY 183 TTTTCAATGGTAGCTTTGTATACGCCAATGCTGTACCAAGTACACAGTCACACTTGA 242  
Db 61 TTCATCATGGTGTCTGTTTCCGTTACCAATGCTGTACACCGTTACCGTGGAAAC 120

QY 243 ACTGCATGGCAGCAACAAACAGAGTCTCAATGATGTTTCTTAAATGATGGCGCAAGC 302  
Db 121 ACCGCTCTGGC--TCAGACACAGTCTCAATGATGTTTCTTAAATGATGGCGCAAGC 177

QY 303 AAGTACCTTAAAGCAAGCATTAGCTTGGACATTCATGACAGTCTCAATTAACAAACT 362  
Db 178 AAGTACCTTAAAGCAAGCATTAGCTTGGACATTCATGACAGTCTCAATTAACAAACT 237

QY 363 TTAGGTACTAGTACAGATTACTCCAGACTCTTTCCTTAAAGCAGGATATTTCTATAGC 422  
Db 238 TTAGGTACTAGTACAGATTACTCCAGACTCTTTCCTTAAAGCAGGATATTTCTATAGC 297

QY 423 AAATTAGATGTTAGGAAGCAGCGTACTGCTAGAGGTACATTTGACTTATGCCAATGTT 482  
Db 298 AAATTAGATGTTAGGAAGCAGCGTACTGCTAGAGGTACATTTGACTTATGCCAATGTT 357

QY 483 GAAGGTAGTACGGTGTAGACATCTTTCCGTTAAATAATCAAAACCCCGCAGGATGGACT 542

Db 358 GAAGGTAGTACGGTGTAGACATCTTTCCGTTAAATAATCAAAACCCCGCAGGATGGACT 417  
QY 543 GGAAACCCCTAATCATGTCAAAATATAAAATGAATGGTTAAATGGTCTATCTTATGTCGGA 602  
Db 418 GGAAACCCCTAATCATGTCAAAATATAAAATGAATGGTTAAATGGTCTATCTTATGTCGGA 477  
QY 603 GATTTCTGGAATGAAGTCAATCTCTATTGCGATAGTCTCGGTGGAGATGCACTCAGAGTC 662  
Db 478 GATTTCTGGAATGAAGTCAATCTCTATTGCGATAGTCTCGGTGGAGATGCACTCAGAGTC 537  
QY 663 AATCCGTTTACAGGAACACCTGATCCCAAAATGTAGGAGTCTGTACCAAAAGCGCGCATG 722  
Db 538 AATCCGTTTACAGGAACACCTGATCCCAAAATGTAGGAGTCTGTACCAAAAGCGCGCATG 597  
QY 723 CGCTATACCGAACAAGAGCTCAAGATGTTAGAGCAATCGTGTATGCTATCTTTAT 782  
Db 598 CGCTATACCGAACAAGAGCTCAAGATGTTAGAGCAATCGTGTATGCTATCTTTAT 657  
QY 783 TATGAAGTCTGCTCAATCTCAACGGCAGAGTGTGATTCGAAGAGTGTCTGTTGATCA 842  
Db 658 TATGAAGTCTGCTCAATCTCAACGGCAGAGTGTGATTCGAAGAGTGTCTGTTGATCA 717  
QY 843 ATGCAATCTTCTGATATACCATCAACGAGAGAGTATTAGTTTACACACAGCTAATGGC 902  
Db 718 ATGCAATCTTCTGATATACCATCAACGAGAGAGTATTAGTTTACACACAGCTAATGGC 777  
QY 903 TACACCATTAACCTACCATCAACGAGTACACCTACTCAAAAATAATATACCAAGGCTAGACCT 962  
Db 778 TACACCATTAACCTACCATCAACGAGTACACCTACTCAAAAATAATATACCAAGGCTAGACCT 837  
QY 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 1022  
Db 838 CTGCTCACTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 897  
QY 1023 CTAGAAAAAGCAATGATTCGCGTCATTGC 1051  
Db 898 CTAGAAAAAGCAATGATTCGCGTCATTGC 926

RESULT 7  
US-08-393-889-10  
Sequence 10, Application US/08393889  
Patent No. 6420152  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
NUMBER OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612

```
REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
US-08-393-889-10

Query Match 18.5%; Score 200; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 8.9e-46;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACGCCCTCTCTTTTCTCTTACTATCTCCTTTAAATTTTCATATTTTAAAAAAC 60
Db 1 GACACGCCCTCTCTTTTCTCTTACTATCTCCTTTAAATTTTCATATTTTAAAAAAC 60
Qy 61 TATTGATAAACTAGTTAAAGCGGTATCTATGTTAGTTCGGAATTTAGAAAGAGG 120
Db 61 TATTGATAAACTAGTTAAAGCGGTATCTATGTTAGTTCGGAATTTAGAAAGAGG 120
Qy 121 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
Db 121 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
Qy 181 AATTTCAATGGTAGCTCTT 200
Db 181 AATTTCAATGGTAGCTCTT 200

RESULT 8
US-08-393-889-10
PCT-US94-09450-10
Sequence 10, Application PC/TUS9409450
GENERAL INFORMATION:
APPLICANT: Beckman Instruments, Inc.
APPLICANT: 2500 Harbor Boulevard
Fullerton, California 92634
TITLE OF INVENTION: Recombinant Dnase B Derived from
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Beckman Instruments, Inc.
STREET: 2500 Harbor Boulevard
CITY: Fullerton
STATE: California
COUNTRY: USA
ZIP: 92634
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/09450
FILING DATE: 18-AUG-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/082,845
FILING DATE: 23-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: May, William H.
REGISTRATION NUMBER: 26,769
REFERENCE/DOCKET NUMBER: 39D-1357 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000

REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
PCT-US94-09450-10

Query Match 18.5%; Score 200; DB 5; Length 200;
Best Local Similarity 100.0%; Pred. No. 8.9e-46;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACACGCCCTCTCTTTTCTCTTACTATCTCCTTTAAATTTTCATATTTTAAAAAAC 60
Db 1 GACACGCCCTCTCTTTTCTCTTACTATCTCCTTTAAATTTTCATATTTTAAAAAAC 60
Qy 61 TATTGATAAACTAGTTAAAGCGGTATCTATGTTAGTTCGGAATTTAGAAAGAGG 120
Db 61 TATTGATAAACTAGTTAAAGCGGTATCTATGTTAGTTCGGAATTTAGAAAGAGG 120
Qy 121 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
Db 121 ACAAGCATATGAATCTACTTGGATCAAGCGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
Qy 181 AATTTCAATGGTAGCTCTT 200
Db 181 AATTTCAATGGTAGCTCTT 200

RESULT 9
US-08-393-889-12
Sequence 12, Application US/08393889
Patent No. 6420152
GENERAL INFORMATION:
APPLICANT: Adams, Craig W.
APPLICANT: Pang, Patty P.-Y.
APPLICANT: Belei, Marina
TITLE OF INVENTION: Recombinant Dnase B Derived from
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheldon & Mak
STREET: 225 South Lake Avenue, Ninth Floor
CITY: Pasadena
STATE: California
COUNTRY: USA
ZIP: 91001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/393,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/082,845
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Michael B.
REGISTRATION NUMBER: 32,612
REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 12:
```





QY 162 AAAAAAT 168  
D<sub>b</sub> 1030 AAAAAAT 1024

RESULT 14

```

US-08-134-557D-1/C
; Sequence 1, Application US/08134557D
; Patent No. 6200802
;
; GENERAL INFORMATION:
;
; APPLICANT: Greene, Marianne E.
; APPLICANT: Blumberg, Bruce
;
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated
;
; TITLE OF INVENTION: Receptor Gamma: Compositions and Methods
;
; NUMBER OF SEQUENCES: 8
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Rokey, Milnamow & Katz, Ltd.
;
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
;
; CITY: Chicago
;
; STATE: IL
;
; COUNTRY: USA
;
; ZIP: 60601
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: Patent In Release #1.0, Version #1.30
;
; APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/134,557D
;
; FILING DATE:

```

Query Match	3.4;	Score 37.2;	DB 3;	Length 1844;
Best Local Similarity	50.6%;	Pred. No. 1.1;		
Matches	90;	Conservative 0;	Mismatches 88;	Indels 0; Gaps 0;
QY	10	TTCTTTTTCCTTACTATCTCTCTTTAAATTTTCATATTTTAAAAAACAATATGTATAA	69	
Db	1844	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGCATTTTCATAATATGTTAAATTTTAA		
QY	70	ACTAGTTAAGTAAGCGTACTATGTGTAGTCGAATTTAGAAAAGGACCAAGCAT	129	
Db	1784	TATTAAGAGTAAATGTGAATGTCTTTATTAACAATATGCAATAAATAGATCATATTC	1725	
QY	130	TGAATCTACTTTGGATCAAGACGGGTTTTTCTTAAAAAATGTGGCTAGTAAATTTTC	187	
Db	1724	TAAACCCCTTTTCTTTTTTAAACCGTTTTTCACAGTAAATTTCTTTATGCTCAGATTTTC	1667	

## RESULT 15

US-09-623-062-1  
; Sequence 1, Application US/09623062  
; Patent No. 6448082  
; GENERAL INFORMATION:



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 14:01:55 ; Search time 3861 Seconds  
(without alignments)  
970.878 Million cell updates/sec

Title: US-08-482-785-7  
Perfect score: 1083  
Sequence: 1 GACACGCCTCTTTTCT.....TGTCGCAAAAGCAAAAGC 1083

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1:	/cgn2_6/ptodata/2/pubnpa/US07_PUBCOMB.seq.*
2:	/cgn2_6/ptodata/2/pubnpa/PCT_NEW_PUB.seq.*
3:	/cgn2_6/ptodata/2/pubnpa/US06_NEW_PUB.seq.*
4:	/cgn2_6/ptodata/2/pubnpa/US06_PUBCOMB.seq.*
5:	/cgn2_6/ptodata/2/pubnpa/US07_NEW_PUB.seq.*
6:	/cgn2_6/ptodata/2/pubnpa/PCTUS_PUBCOMB.seq.*
7:	/cgn2_6/ptodata/2/pubnpa/US08_NEW_PUB.seq.*
8:	/cgn2_6/ptodata/2/pubnpa/US08_PUBCOMB.seq.*
9:	/cgn2_6/ptodata/2/pubnpa/US09A_PUBCOMB.seq.*
10:	/cgn2_6/ptodata/2/pubnpa/US09B_PUBCOMB.seq.*
11:	/cgn2_6/ptodata/2/pubnpa/US09C_PUBCOMB.seq.*
12:	/cgn2_6/ptodata/2/pubnpa/US09_NEW_PUB.seq.*
13:	/cgn2_6/ptodata/2/pubnpa/US09_NEW_PUB.seq2.*
14:	/cgn2_6/ptodata/2/pubnpa/US10A_PUBCOMB.seq.*
15:	/cgn2_6/ptodata/2/pubnpa/US10B_PUBCOMB.seq.*
16:	/cgn2_6/ptodata/2/pubnpa/US10_NEW_PUB.seq.*
17:	/cgn2_6/ptodata/2/pubnpa/US60_NEW_PUB.seq.*
18:	/cgn2_6/ptodata/2/pubnpa/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	9	US-09-119-900-7
2	922.6	85.2	940	9	US-09-119-900-11
3	853.8	78.8	937	9	US-09-119-900-14
4	200	18.5	200	9	US-09-119-900-10
5	106.6	9.8	182	9	US-09-119-900-12
C 6	46.2	4.3	6106	13	US-10-311-455-1445
C 7	45	4.2	3673778	13	US-10-312-841-2
C 8	43.6	4.0	744802	12	US-10-292-798-1369
C 9	43.4	4.0	8588	13	US-10-240-453-199
C 10	43.4	4.0	8588	15	US-10-239-676-177
C 11	43.4	4.0	16373	13	US-10-311-455-591
C 12	43.2	4.0	9770	13	US-10-311-455-5
C 13	42.6	3.9	13449	13	US-10-311-455-1357
C 14	42.4	3.9	7503	13	US-10-311-455-1521
C 15	41	3.8	5678	13	US-10-311-455-1112

C	16	40.8	3.8	1049	9	US-09-800-729-67
C 17	40.8	3.8	6101	13	US-10-311-455-1447	
C 18	40.8	3.8	8238	13	US-10-311-455-1961	
C 19	40.6	3.7	3673778	13	US-10-312-841-1	
C 20	40.2	3.7	6210	13	US-10-240-453-346	
C 21	40.2	3.7	37973	13	US-10-311-455-2170	
C 22	40	3.7	6208	13	US-10-311-455-168	
C 23	40	3.7	6815	13	US-10-311-455-644	
C 24	40	3.7	6815	13	US-10-240-453-50	
C 25	40	3.7	6815	15	US-10-239-676-50	
C 26	40	3.7	19082	13	US-10-311-455-599	
C 27	39.8	3.7	5662	15	US-10-198-846-13759	
C 28	39.8	3.7	6114	13	US-10-311-455-734	
C 29	39.8	3.7	17421	13	US-10-240-453-55	
C 30	39.8	3.7	17421	15	US-10-239-676-53	
C 31	39.6	3.7	5415	13	US-10-311-455-618	
C 32	39.6	3.7	6948	13	US-10-311-455-1382	
C 33	39.6	3.7	12763	13	US-10-311-455-275	
C 34	39.6	3.7	3673778	13	US-10-312-841-2	
C 35	39.4	3.6	7057	13	US-10-311-455-1822	
C 36	39.4	3.6	7057	13	US-10-240-485-148	
C 37	39.4	3.6	10945	13	US-10-240-453-228	
C 38	39.4	3.6	40862	13	US-10-311-455-2045	
C 39	39.2	3.6	15592	13	US-10-311-455-1299	
C 40	39	3.6	5572	13	US-10-311-455-1398	
C 41	39	3.6	5572	13	US-10-240-452-60	
C 42	39	3.6	14708	13	US-10-311-455-2217	
C 43	39	3.6	14708	13	US-10-240-453-323	
C 44	39	3.6	14708	15	US-10-239-676-221	
C 45	39	3.6	15416	13	US-10-311-455-2204	

#### ALIGNMENTS

#### RESULT 1

US-09-119-900-7  
; Sequence 7, Application US/09119900  
; Patent No. US20020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 7:

```
SEQUENCE CHARACTERISTICS:
LENGTH: 1083 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
FEATURE:
NAME/KEY: CDS
LOCATION: 129..944
US-09-119-900-7

Query Match      100.0%; Score 1083; DB 9; Length 1083;
Best Local Similarity 100.0%; Pred. No. 9e-266;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACGCCCTCTCTTTTCTCCTTACTATCTCTCTTTAAATTTTCATATTTTAAATAAAC 60
DB 1 GACACGCCCTCTCTTTTCTCCTTACTATCTCTCTTTAAATTTTCATATTTTAAATAAAC 60

QY 61 TATTGATAAAGTAACTAGTAAAGCGGTATCTAGTGTAGTGGAAATTTAGAAAAGAGG 120
DB 61 TATTGATAAAGTAACTAGTAAAGCGGTATCTAGTGTAGTGGAAATTTAGAAAAGAGG 120

QY 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAATAATGTCGGGTAGTAA 180
DB 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAATAATGTCGGGTAGTAA 180

QY 181 AATTTTCAATGGTAGCTCTTTGATCAGCCACAATGGCTGTAACACAGTCACACTTGAAA 240
DB 181 AATTTTCAATGGTAGCTCTTTGATCAGCCACAATGGCTGTAACACAGTCACACTTGAAA 240

QY 241 ATACTGCACTGCGACGACAAACACAGGTCTCAATGATGTTGTTCTAAATGATCGGCCAA 300
DB 241 ATACTGCACTGCGACGACAAACACAGGTCTCAATGATGTTGTTCTAAATGATCGGCCAA 300

QY 301 GCAAGTACTATAACGAAACATTAGCTTGGACATTCATGACAGTCTCAATCTATTACAAA 360
DB 301 GCAAGTACTATAACGAAACATTAGCTTGGACATTCATGACAGTCTCAATCTATTACAAA 360

QY 361 CTTTAGGTACTAGTCAGATTAATCTCAGACCTCTTCTTAAAGCAGGAGATTTCTCTATA 420
DB 361 CTTTAGGTACTAGTCAGATTAATCTCAGACCTCTTCTTAAAGCAGGAGATTTCTCTATA 420

QY 421 GCAATTTAGATGAGTTAGGAGGACGGGTACTGCTAGAGGTACATTGACTTATGCCAATG 480
DB 421 GCAATTTAGATGAGTTAGGAGGACGGGTACTGCTAGAGGTACATTGACTTATGCCAATG 480

QY 481 TTGAAGGTAGCTACGGTGTAGACAAATCTTTTCGGTAAATCAAAACCCCGCAGATGGA 540
DB 481 TTGAAGGTAGCTACGGTGTAGACAAATCTTTTCGGTAAATCAAAACCCCGCAGATGGA 540

QY 541 CTGGAACCCCTAATCATGTCAAATATAAATGAATGTTAAATGGTCTATCTTATGTCG 600
DB 541 CTGGAACCCCTAATCATGTCAAATATAAATGAATGTTAAATGGTCTATCTTATGTCG 600

QY 601 GAGATTTCTGGAATAGAGTCAATCTCAGATGTCGGTGGAGATGCCACTCAGAG 660
DB 601 GAGATTTCTGGAATAGAGTCAATCTCAGATGTCGGTGGAGATGCCACTCAGAG 660

QY 661 TCAATGCGGTTPACAGGAACAGTACCCAAATGTTAGGAGGTCTGTACCAAAAAGCGGCA 720
DB 661 TCAATGCGGTTPACAGGAACAGTACCCAAATGTTAGGAGGTCTGTACCAAAAAGCGGCA 720

QY 721 TGGCCTATACCGAACAGAGCTCAAGATGTTTAGAAGCAATCGTGATGCGTATCTTT 780
DB 721 TGGCCTATACCGAACAGAGCTCAAGATGTTTAGAAGCAATCGTGATGCGTATCTTT 780

QY 781 ATTATGAAGTCGCTCCCAATCTACAAACGACAGGTGATTTCCAAAGAGCTGTCGTAT 840
DB 781 ATTATGAAGTCGCTCCCAATCTACAAACGACAGGTGATTTCCAAAGAGCTGTCGTAT 840

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1083 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 129..944
; US-09-119-900-7

781 ATTATGAAGTCGCTCCCAATCTACAAACGACAGGTGATTTCCAAAGAGCTGTCGTAT 840
841 CAATGCAATCTTCTGATAATACCAATCAACGAGAAAGTATTAGTTTACAACACAGCTAATG 900
841 CAATGCAATCTTCTGATAATACCAATCAACGAGAAAGTATTAGTTTACAACACAGCTAATG 900
901 GCTACACATTAACTACTACCAATACCGGTACACTACTCAAAAATAATACCAAAAGGCTAGAC 960
901 GCTACACATTAACTACTACCAATACCGGTACACTACTCAAAAATAATACCAAAAGGCTAGAC 960
961 CTCTGCTCACTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAA 1020
961 CTCTGCTCACTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAA 1020
1021 TACTAGAAAAAGCAATGATTCGCGTCAATTCCTTTTATGAATTTGTCAAAAAGCAAAA 1080
1021 TACTAGAAAAAGCAATGATTCGCGTCAATTCCTTTTATGAATTTGTCAAAAAGCAAAA 1080
1081 AGC 1083
1081 AGC 1083

RESULT 2
US-09-119-900-11
; Sequence 11, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Bellei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 940 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; US-09-119-900-11
```

Query Match 85.2%; Score 922.6; DB 9; Length 940;  
Best Local Similarity 99.6%; Pred. No. 6.8e-225;  
Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAATAATGTCGGCTAGTAAAA 182  
Db 1 ATGGATCCGAATCTACTTGGATCAAGACGGGTTTTTCTAAATAATGTCGGCTAGTAAAA 60

Qy 183 TTTTCAATGCTAGCTTGTATGACCAATGCTGTAAACAAGTCACACTTGAAAT 242  
Db 61 TTTTCAATGCTAGCTTGTATGACCAATGCTGTAAACAAGTCACACTTGAAAT 120

Qy 243 ACTGCACTGCGACGACAAACACAGGCTCTCAATGATGTTCTTAATGATGCGCAAGC 302  
Db 121 ACTGCACTGCGACGACAAACACAGGCTCTCAATGATGTTCTTAATGATGCGCAAGC 180

Qy 303 AAGTACCTAAACGAAGCAATTAGCTTGGACATCTCAATGACAGTCTCTAACTATTACAAACT 362  
Db 181 AAGTACCTAAACGAAGCAATTAGCTTGGACATCTCAATGACAGTCTCTAACTATTACAAACT 240

Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGAGATATTTCTATAGC 422  
Db 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGAGATATTTCTATAGC 300

Qy 423 AAATTAGATGAGTTAGGAAGGACGCGTACTGCTAGAGGTACATGACTTATGCAATGTT 482  
Db 301 AAATTAGATGAGTTAGGAAGGACGCGTACTGCTAGAGGTACATGACTTATGCAATGTT 360

Qy 483 GAAGGTAGTACGGTGTAGACAACTCTTCGGTAAATAATCAAAACCCGCGAGATGACT 542  
Db 361 GAAGGTAGTACGGTGTAGACAACTCTTCGGTAAATAATCAAAACCCGCGAGATGACT 420

Qy 543 GGAACCCCTAATCATGTCAATAATAAATTAATGTTAAATGTTCTTATGTCGGA 602  
Db 421 GGAACCCCTAATCATGTCAATAATAAATTAATGTTAAATGTTCTTATGTCGGA 480

Qy 603 GATTTCTGGAATGAGTATCTCATGACAGATGCTCGGTGGAGATGCACTCAGATC 662  
Db 481 GATTTCTGGAATGAGTATCTCATGACAGATGCTCGGTGGAGATGCACTCAGATC 540

Qy 663 AATGCGGTACAGGAACAGCTACCCAAATGTAGGAGGTCTGACCAAAAGGCGCATG 722  
Db 541 AATGCGGTACAGGAACAGCTACCCAAATGTAGGAGGTCTGACCAAAAGGCGCATG 600

Qy 723 CGCTATACCGAACAAGAGCTCAAGAATGTTAGAGCAAAATCGTGATGCTATCTTTAT 782  
Db 601 CGCTATACCGAACAAGAGCTCAAGAATGTTAGAGCAAAATCGTGATGCTATCTTTAT 660

Qy 783 TATGAAGTCGCTCAATCTACACGCGAGAGTGTGATTCGAAGAGCTGTCGGTATCA 842  
Db 661 TATGAAGTCGCTCAATCTACACGCGAGAGTGTGATTCGAAGAGCTGTCGGTATCA 720

Qy 843 ATGCAATCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGCG 902  
Db 721 ATGCAATCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGCG 780

Qy 903 TACACCAATTAACCTACCAATACCGGTACACCTTACTCAAAAATAATACCAAAAGGCTAGACCT 962  
Db 781 TACACCAATTAACCTACCAATACCGGTACACCTTACTCAAAAATAATACCAAAAGGCTAGACCT 840

Qy 963 CTGCTCAGTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 1022  
Db 841 CTGCTCAGTAGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAATA 900

Qy 1023 CTAGAAAAGCAATGATGCGGTATGTC 1051  
Db 901 CTAGAAAAGCAATGATGCGGTATGTC 929

RESULT 3

US-09-119-900-14

; Sequence 14, Application US/09119900

; Patent No. US20020081622A1

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: US/09/119,900  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 937 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..819  
US-09-119-900-14

Query Match 78.8%; Score 853.8; DB 9; Length 937;  
Best Local Similarity 95.7%; Pred. No. 2.4e-207;  
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAATAATGTCGGCTAGTAAAA 182  
Db 1 ATGGATCCGAATCTACTTGGATCAAGACGGGTTTTTCTAAATAATGTCGGCTAGTAAAA 60

Qy 183 TTTTCAATGCTAGCTTGTATGACCAATGCTGTAAACAAGTCACACTTGAAAT 242  
Db 61 TTTTCAATGCTAGCTTGTATGACCAATGCTGTAAACAAGTCACACTTGAAAT 120

Qy 243 ACTGCACTGCGACGACAAACACAGGCTCTCAATGATGTTCTTAATGATGCGCAAGC 302  
Db 121 ACCGCTCTGCG--TCAGACACAGGCTCTCAATGATGTTCTTAATGATGCGCAAGC 177

Qy 303 AAGTACCTAAACGAAGCAATTAGCTTGGACATCTCAATGACAGTCTTAATGATGCGCAAGC 362  
Db 178 AAGTACCTAAACGAAGCAATTAGCTTGGACATCTCAATGACAGTCTTAATGATGCGCAAGC 237

Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGAGATATTTCTATAGC 422  
Db 238 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCCTAAAGCAGAGATATTTCTATAGC 297



Qy	336	AATGACAGTCTCTAACTATTACAAACT	362
Db	5050	AAAAATATCTCTAACTTTTAAAAATCT	5024

	Query Match	4.2%;	Score 45;	DB 13;	Length 3673778;
	Best Local Similarity	49.8%;	Pred. No. 44;		
	Matches 114;	Conservative 0;	Mismatches 115;	Indels 0;	Gaps 0;
QY	851	TTCTGATAATACCATCAACGAGAGAAGCTATTAGTGTTTACAACACAGACTAAATGGCTACACCAT	910		
DB	470019	TTAAANATTATCACAAAATTATTAACAACAACCTATAACATATAATACTATTAAAAAAAACCT	46999		
QY	911	TAACTACCATTAACGGTACACCTACTCAAAAATAATACCAAAAGGCTAGACCTCTGCTCAC	970		
DB	469959	TAAAAACAATTATTATTAAACTACTTCAAAAATATATACAAAATTTTAAATACCATATTTCAC	46999		
QY	971	TAGGCCCTAGCTTTTTTACATCAAAAAAAGCAATGACCTATAGAAGTAAAAAATACTAGAAAA	1030		
DB	469899	CAACTATTACATTTTACACAAAAATAAATTTTCTATATATTCTTTTAATTAATAAAAAAA	4698		
QY	1031	AGCAATGATTGGCGTCATTCGCTTTTTTATGAAATTTGTGCAAAAAGCAAAA	1079		
DB	469839	AACAATAATTTCTATTATTACTAAATATATATTCTACATATAAATTTTAA	469791		

US-10-292-798-1369/c  
; Sequence 1369, Application US/10292798  
; Publication No. US2003023583A1  
; GENERAL INFORMATION:

APPLICANT: SUWA, MAKIKO  
APPLICANT: ASAI, KIYOSHI  
APPLICANT: AKIYAMA, YUTAKA  
APPLICANT: ABURATANI, HIROYUKI  
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
FILE REFERENCE: 084335/166  
CURRENT APPLICATION NUMBER: US/10/292,798  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: 10/017,161  
PRIOR FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: JP 2001-246789  
NUMBER OF SEQ ID NOS: 2070  
SOFTWARE: Patentin ver. 2.1  
SEQ ID NO 1369  
LENGTH: 744802  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
LOCATION: source  
FEATURE:  
LOCATION: (1)..(744802)  
NAME/KEY: CDS  
LOCATION: (201)..(246)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (25640)..(25677)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (27078)..(27094)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (141192)..(141769)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (159571)..(159606)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (174525)..(174575)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (234891)..(235013)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (235514)..(235560)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (279677)..(279729)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (408660)..(409123)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (409204)..(409669)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (428381)..(428396)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (472204)..(472330)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (714252)..(714355)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (714447)..(714529)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (739794)..(739891)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (744484)..(744602)

FEATURE:  
NAME/KEY: modified base  
LOCATION: (51812)..(51911)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (57122)..(57221)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (79368)..(79467)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (293951)..(294050)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (310089)..(310188)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (332935)..(332935)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (332992)..(332992)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (362002)..(362101)  
OTHER INFORMATION: a, t, c, g, unknown or other  
FEATURE:  
NAME/KEY: modified base  
LOCATION: (639781)..(639880)  
OTHER INFORMATION: a, t, c, g, unknown or other  
US-10-292-798-1369

Query Match 4.0%; Score 43.6; DB 12; Length 744802;  
Best Local Similarity 56.2%; Pred. NO. 43;  
Matches 82; Conservative 0; Mismatches 64; Indels 0; Gaps 0;  
QY 46 ATTTTAAAGAACTATTGATAAACTAGTTAAAGTAGCGGTATCTAGTTAGTTCGCG 105  
Db 56341 ACTATTACAAAGCACCTTCAGATAATAGAAAACAAAGGAACTATATGACAACT 56282  
QY 106 AAATTAGAAAAGAGACAGCATATGATCTCTTGGATCAAGCGGGTTTTTCTAAAA 165  
Db 56281 GAATTAAACAAATATATAAAAGATAACATGTTGACCAAGTTGGTTTATCTTAAA 56222  
QY 166 AATGCGGCTAGTAAATTTTCAATG 191  
Db 56221 ATGCAGGTGAGTTTAAATTTTGAAG 56196

RESULT 9  
US-10-240-453-199/c  
Sequence 199, Application US/10240453  
Publication No. US20030148326A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA  
TITLE OF INVENTION: Transcription  
TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated  
TITLE OF INVENTION: with DNA Transcription  
FILE REFERENCE: 5013.1009  
CURRENT APPLICATION NUMBER: US/10/240,453  
CURRENT FILING DATE: 2002-10-02  
PRIOR APPLICATION NUMBER: PCT/EP01/03973  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: DE 10019058.8

PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 350  
SEQ ID NO 199  
LENGTH: 8588  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4416, 4418, 4430, 4434..4435, 5243, 5245, 5612)  
US-10-240-453-199

Query Match 4.0%; Score 43.4; DB 13; Length 8588;  
Best Local Similarity 51.9%; Pred. No. 4.1;  
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 895 CTAATGCTACACCATTAACCTACCAACGCTACACCTCTCAAAAATAATACCAAGG 954  
DB 7748 CTATTATCTACAACATAATTAACAACCAACCAAACTCTTAATATATATACG 7689  
QY 955 CTAGACCTCTGCTCAGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAG 1014  
DB 7688 CTTAACTCTACAAAATAACTTAACCTTCTACACTCAACCATCAATTTTCAACCC 7629  
QY 1015 TAAATACTAGAAAAAGCAATGATCCGCTCATGCTTTTATGAATTTGTGCAAAAG 1074  
DB 7628 TCAATAAATTAATAAATTAATAAATAAACTACAAACATATTTACTAATATATATCTAAA 7569  
QY 1075 CAAAAAGC 1083  
DB 7568 AAAAAAAC 7560

RESULT 10  
US-10-239-676-177/c  
Sequence 177, Application US/10239676  
Publication No. US20030082609A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation  
FILE REFERENCE: 5013.1003  
CURRENT APPLICATION NUMBER: US/10/239,676  
PRIOR FILING DATE: 2002-09-24  
PRIOR APPLICATION NUMBER: PCT/EP01/03968  
DE 10019058.8  
DE 10019173.8  
DE 10032529.7  
DE 10043826.1  
PRIOR FILING DATE: 2001-04-06  
2000-04-06  
2000-04-07  
2000-06-30  
2000-09-01  
NUMBER OF SEQ ID NOS: 228  
SEQ ID NO 177  
LENGTH: 8588  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (4416, 4418, 4430, 4434..4435, 5243, 5245, 5612)  
US-10-239-676-177

Query Match 4.0%; Score 43.4; DB 15; Length 8588;  
Best Local Similarity 51.9%; Pred. No. 4.1;  
Matches 98; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
QY 895 CTAATGCTACACCATTAACCTACCAACGCTACACCTCTCAAAAATAATACCAAGG 954  
DB 7748 CTATTATCTACAACATAATTAACAACCAACCAAACTCTTAATATATATACG 7689  
QY 955 CTAGACCTCTGCTCAGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAG 1014  
DB 7688 CTTAACTCTACAAAATAACTTAACCTTCTACACTCAACCATCAATTTTCAACCC 7629  
QY 1015 TAAATACTAGAAAAAGCAATGATCCGCTCATGCTTTTATGAATTTGTGCAAAAG 1074  
DB 7628 TCAATAAATTAATAAATTAATAAATAAACTACAAACATATTTACTAATATATATCTAAA 7569  
QY 1075 CAAAAAGC 1083  
DB 7568 AAAAAAAC 7560

RESULT 11  
US-10-311-455-591/c  
Sequence 591, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 591  
LENGTH: 16373  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-591

Query Match 4.0%; Score 43.4; DB 13; Length 16373;  
Best Local Similarity 50.2%; Pred. No. 5.8;  
Matches 107; Conservative 0; Mismatches 106; Indels 0; Gaps 0;  
QY 819 ATTCCAAGAGCTGCTGGTATCAATGCAATCTTCTGATTAATACCATCAACGAGAAAGTA 878  
DB 4652 ATATAATAAATATCTAAATATAATCAAAAACCTCTCATAAAAAATCAACAATTAATC 4593  
QY 879 TTAGTTTACACACAGCTAATGCTACACCATTAACCTACCATCAACGCTACCTACTCAA 938  
DB 4592 TTAACTTAAAAAATCAAAAAACCAACCATCATCATATATCAATATATAATCAACTTAAA 4533  
QY 939 AAATAATACCAAAAGGCTAGACCTCTGCTCAGCTAGGCTAGCTTTTACATCAAAAAAG 998  
DB 4532 CAATAAAACACACGACCAACCAAAAAAATCAACATATAAATCAAAAAAATC 4473  
QY 999 CAATGACTATAGAAAGTAAAAATCTAGAAAA 1031  
DB 4472 TAATAAACTAAAAACCAATATAAAACAAAAATAA 4440

RESULT 12  
US-10-311-455-5/c

```
; Sequence 5, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 5
; LENGTH: 9770
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-5

Query Match
Best Local Similarity 4.0%; Score 43.2; DB 13; Length 9770;
Matches 99; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 849 TCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGGCTACACC 908
DB TTTCTTAACTACAAACCAAAAAACCAATACTACTACACAAAAACCCCTAAAAA 3988
QY 909 ATTAATACATACCGGTACACTCTCAAAAATAATACCAAAAGGCTAGACCTCTGCTC 968
DB ACTAAATTCACACCGCAACCAACAAACAAATTTACCTTTCACTTCGCCCCAC 3928
QY 969 ACTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAAAATACTAGAA 1028
DB TCATCTCTAAATTCMAAAAAAACAACCACTTAAATACTAAAAAATAAAAAA 3868
QY 1029 AAGCAATGATT 1040
DB AAAAAAACGATT 3856

RESULT 13
US-10-311-455-1357/c
; Sequence 1357, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1357
; LENGTH: 13449
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
```

US-10-311-455-1357

```
Query Match
Best Local Similarity 3.9%; Score 42.6; DB 13; Length 13449;
Matches 81; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 887 CAACACAGCTAATGGGTACACCAATTAATACCAATACGCTACACTCTCAAAAATAATA 946
DB CAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11824
QY 947 CAAAAGGCTAGACCTCTGCTCACTAGGCGCTAGCTTTTACATCAAAAAAGCAATGACT 1006
DB AAAAATTTCTAACTTAAACCTTAAACCAATCTTTTAAATAATCCAAACCAAAAAAATAA 11764
QY 1007 ATAGAAAGTAAAAATACTAGAAAAA 1031
DB AAAAAATAAATAAATAAATAAATAAATAA 11739
```

RESULT 14

```
US-10-311-455-1521/c
; Sequence 1521, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1521
; LENGTH: 7503
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1521
```

```
Query Match
Best Local Similarity 3.9%; Score 42.4; DB 13; Length 7503;
Matches 91; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 864 ATCAACGAGAAAGTATTAGTTTACACACAGCTAATGGCTACACCAATTAATACCTAAC 923
DB AACTACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6091
QY 924 GGTACACCTACTCAAAAATAATACCAAAAGGCTAGACCTCTGCTCACTAGCGCTAGCTTT 983
DB GTTTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 6031
QY 984 TTACATCAAAAAGCAATGACTATAGAAAGTAAAAATACTAGAAAAAGCAA 1035
DB TTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5979
```

RESULT 15

```
US-10-311-455-1112/c
; Sequence 1112, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
```



Search completed: January 5, 2004, 18:35:40  
Job time : 3874 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:53:50 ; Search time 2237 Seconds  
(without alignments)  
11766.539 Million cell updates/sec

Title: US-08-482-785-7

Perfect score: 1083

Sequence: 1 GACACGCCCTCTTTTCT.....TGTCAAAAGCAAAAGC 1083

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vit:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gsal:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50.2	4.6	1201	13	BX402521
c 2	49.8	4.6	1146	13	BX462546
3	48.6	4.5	1201	13	BX456567
c 4	48.2	4.5	427	13	BX403499

c 5	47.4	4.4	627	28	AZ522909
c 6	47.2	4.4	254	12	BI743690
c 7	47.2	4.4	712	13	BI743690
c 8	46.8	4.3	878	29	CNS0187R
c 9	46.8	4.3	881	28	AZ691287
10	46.6	4.3	1101	29	CNS00E0D
11	46.4	4.3	885	28	BH132872
12	46.4	4.3	1201	13	BX344084
13	46.2	4.3	1201	9	AL514423
14	46.2	4.3	1201	13	BX356851
15	45.8	4.2	1124	13	BX436282
16	45.4	4.2	1037	13	BX359859
17	45.2	4.2	882	28	BH161458
18	45.2	4.2	890	28	AZ545520
19	45	4.2	739	28	BH604544
c 20	45	4.2	880	28	AZ550450
c 21	45	4.2	1201	13	BX376097
c 22	45	4.2	1201	13	BX445758
23	44.8	4.1	882	28	AZ677620
24	44.6	4.1	1101	29	CNS002FU
25	44.6	4.1	1201	13	BX422711
26	44	4.1	865	28	AZ677118
27	44	4.1	895	28	BH158314
28	44	4.1	911	28	AZ533976
29	44	4.1	954	28	BH137041
c 30	43.8	4.0	742	29	CC082163
c 31	43.8	4.0	1200	13	BX437758
32	43.8	4.0	1521	12	BM913345
33	43.6	4.0	670	29	BX175553
34	43.6	4.0	689	28	BH549374
35	43.6	4.0	712	29	BX227138
36	43.6	4.0	834	29	BZ486924
37	43.6	4.0	843	28	AZ682679
38	43.6	4.0	895	28	BH164351
39	43.6	4.0	907	28	BH134191
40	43.6	4.0	913	28	BH157356
41	43.6	4.0	917	28	AZ535994
c 42	43.6	4.0	1101	29	CNS00HC2
43	43.6	4.0	1201	13	BX446296
44	43.4	4.0	655	29	BZ504420
c 45	43.4	4.0	850	28	AZ551139

#### ALIGNMENTS

RESULT 1  
BX402521  
LOCUS  
DEFINITION  
BX402521 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
clone CSOD1046YB22 5-PRIME, mRNA sequence.  
1201 bp mRNA linear EST 13-MAY-2003  
EST.  
BX402521 GI:30622835  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1201)  
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CSI10122D12QPl.  
Location/Qualifiers  
1. .1201  
/organism="Homo sapiens"

```

/mot_type="mrna"
/db_xref="taxon:9606"
/clone="CS0D1046Y822"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT      457 a   187 c   132 g   316 t   109 others
ORIGIN
Query Match      4.6%; Score 50.2; DB 13; Length 1201;
Best Local Similarity 35.3%; Pred. No. 40;
Matches 174; Conservative 59; Mismatches 260; Indels 0; Gaps 0;

QY 14 TTTTCTCCTTACTACTCTCTTAATTTTCATATTTTAAATAAACTATGTGATAACTA 73
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
702 KTTTYYTTTAAAWTTTKATTAATAATTTTWTATKTAATAAAAWTTTYYTATAAAA 761
QY 74 GTTAAGTAAGCGTATACATGCTAGTTCGCAATATGAAAGAGGACAGCATATGAA 133
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
762 AWTTTTATATTTTATTTTADTKKKGAGKWTAAATTKAGRRATTTTAAWTTTAA 821
QY 134 TCTACTTGATCAAGACGGGTTTTTCTAAATAATGTCGGCTAGTAAATAATTTCAATGGT 193
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
822 TWAATTTATKTAATAATTTTWTWWDGGGTTTAAATAAAATTTTAAAVGTK 881
QY 194 AGCTTTGTATCAGGCACATGGCTGTAAACAGTCACCTTGAAATATCTGCATGGC 253
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
882 AAATTAWAGGAATATATATWGTAKAAATTAATTAATAAAAWAAATAAAGATGTAT 941
QY 254 ACGCAACACACAGGCTCAATATGATGTGTTCTTAATGATCGCGCAAGTACCTAAA 313
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
942 AAAAAAATAAAGGATATTTTCWTTTATATAAAGAGAGAAAAATTTTATATA 1001
QY 314 CGAAGCATTAGCTGGACATTCATGACAGCTCACTTAACTATTACAAAATTTAGGTACTAG 373
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
1002 TAAATAAATTTTAAAWATWTAATTKSADAACTAATTAATTWAGATAAAATTTTK 1061
QY 374 TCAGATTACTCCAGCAGCTCTTCTTAAAGCAGGAGATATCTCTATAGCAATTAGATGA 433
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
1062 WAAAAAATAWAAAAAATAATGTGCGAAADAATTTWAAAAAARARARADAAAAAGR 1121
QY 434 GTTAGGAGGAGCGCTACTGCTAGTAGTACATTTGACTTATCCAAATTTGAAGGTAGCTA 493
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
1122 GKTAGAAAAAATAAAGGAAGAAWAAWAAAAAAYTTKORRAAAAAAAGRAAGCAAGA 1181
QY 494 CGGTGTTAGACAA 506
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :
1182 AGGGRGAAAAAA 1194

RESULT 2
BX462546/c
LOCUS
DEFINITION
BX462546 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
clone CS0DH003YP10 3-PRIME, mRNA sequence.
ACCESSION
BX462546
VERSION
BX462546.1 GI:31027451
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1146)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

REFERENCE
AUTHORS
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE
Full-length cDNA libraries and normalization
JOURNAL
Unpublished
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 994.f For more information about this cluster, see <http://www.genoscope.cns.fr/> <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DH003DH05NP1&cluster=994.f>. Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DH003DH05NP1.

#### FEATURES

Location/Qualifiers  
1..1146  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="CS0DH003YP10"  
/tissue\_type="T CELLS (JURKAT CELL LINE)"  
/cell\_line="JURKAT CELL LINE"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE)"  
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
BASE COUNT 216 a 127 c 154 g 436 t 213 others  
ORIGIN

Query Match 4.6%; Score 49.8; DB 13; Length 1146;  
Best Local Similarity 41.2%; Pred. No. 48;  
Matches 73; Conservative 34; Mismatches 70; Indels 0; Gaps 0;  
QY 8 CCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAATAAACTATTTGAT 67  
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :  
1116 YCTTYYTTTYYTAAATTAATTTTAAATTTTAAATTTTAAATAAAATTTTAAATAA 1057  
QY 68 AAATCTAGTTAAGTAAAGCTATCTAGTGTAGTTAGCGAAATTTAGAAAAGAGCAAGCA 127  
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :  
1056 TWWTAAWAAATTCAAAAATTTWTAATTTTAAATAAAATTTAAATAAAATTTAAATAA 997  
QY 128 TATGAATCTACTGGATCAAGCGGCTTTTCTTAAATAATGTCGCTAGTAAATTT 184  
Db :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :|||: :  
996 WATTTTAAATTTTAAATTTTAAATTTTAAATAAAATTTTAAATAAAATTTTAAATAA 940

RESULT 3  
BX456567 1201 bp mRNA linear EST 22-MAY-2003  
LOCUS  
DEFINITION  
BX456567 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP002YN23  
3-PRIME, mRNA sequence.

ACCESSION  
BX456567  
VERSION  
BX456567.1 GI:31024832  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
1 (bases 1 to 1201)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE  
Full-length cDNA libraries and normalization  
JOURNAL  
Unpublished  
COMMENT  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0CAP002CGI2NP1.

#### FEATURES

Location/Qualifiers  
1..1201  
/organism="Homo sapiens"  
/mol\_type="mrna"  
/db\_xref="taxon:9606"  
/clone="CS0CAP002YN23"  
/tissue\_type="THYMUS"

```
/clone lib="Homo sapiens THYMUS"
/notes=vector: pcWSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pcWSPORT 6 vector.
Library was not normalized "
```

BASE COUNT	
ORIGIN	
457 a	156 c 105 g 352 t 131 others

Query Match 4.5%; Score 48.6; DB 13; Length 1201;  
Best Local Similarity 38.8%; Pred. No. 76;  
Matches 116; Conservative 38; Mismatches 145; Indels 0; Gaps 0;

Qy	16	TTTCTCCTTACTACTCTCTCTTTTAAATTTTTCATATTTTAAATAAACTATTGTATAAACTAGT	75
Db	884	TTTTTTTTTMCACAAATTTTTTTTTTTTHAMMTWTTTTTTAAATAAAATTTTTTTTTTAAATAAAAT	943

Qy	76	TAAGTAACGGTATACTATGTTAGTTTAGCGAAATTAGAAAAAGAGGCAAGCATATGAATC	135
<hr/>			
Db	944	TWAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTKTWAAAAAAAAAAAAAAAAAWTTTTTTTA	1003

Qy 136 TACTTGATCAAGACGGTTTTTTTCTTAAAAAATGTCGGCTAGTAAAAATTTTCAATGGTAG 195

Db 1004 AAAAATAAAAAATTTTTTTTTTTTTTTTTTTTAAAAAAATWTAAAAAAATTTTAAWTAAAAA 1063

	Qy	196	CTCTTGTTATCAGCCACAAATGGCTGTAAACAACAGTCACACTTGAATAAATTCTGCACCTGGCAC	255
	Dβ	1064	AAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAACCCCCTDKAAAAAAAWKAKKKCCCC ::	1123

**OY**

256 GACAAACACAGGTCTCAAATGATGTTTCTTAATGATGGCGAAGCAAGTAGCTACTAAC 314  
::: :::  
**D8** 1124 CMAATTCTTKKKAATAKTGWABBCBCCAIAAAAAADNAAAAMWMDPFAAAARCC 1182  
::: :::

RESULT 4  
RX403499/C

LOCUS	BX403499	427 bp	linear	EST 13-MAY-2000
DEFINITION	BX403499	Homo sapiens PLACENTA	Homo sapiens	cDNA clone CLO8A002ZH05
ACCSSION	BX403499	3-PRIME	mRNA sequence.	

VERSION	BOX03499.1	GI:30635017
KEYWORDS	EST.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE  
1. (bases 1 to 427)  
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

**TITLE** Full-length cDNA libraries and normalization  
**COMMENT** Unpublished  
**CONTACT** Genoscope

BP 191 31006 EVRY cedex - France  
Email: [segre@genoscope.cns.fr](mailto:segre@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of

invitrogen; contact : peng liang email : [liangpeng@invitrogen.com](mailto:liangpeng@invitrogen.com) ORL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CL0BA002ZH05FP1.  
Location/Qualifiers

```

1.42 /
/organism="Homo sapiens"
/mol_type="mrna"
/db xref="taxon:9606"
source

```

```

/clone="CLOB002ZH05"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens
PLACENTA"
/notes="Vector: PCMVSPORT
6: 1st strand
CDNA was primed

```

with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

	BASE COUNT	248 a	16 c	16 g	59 t	88 others
ORIGIN						

Query Match 4.5%; Score 48.2; DB 13; Length 427;  
Best Local Similarity 32.3%; Pred. No. 1.2e+02;  
Matches 59; Conservative 51; Mismatches 73; Indels 0; Gaps 0;

Qy	10	TCTTTTTTCTCCTTACTATCTCCITTAATTTCATATTTTAAAAAACTATTGATAA	69
Db	229	TTTITTTTTTTTTTTTTTTNAAAANNNTTTTTTTTTTTTTTTTAAAAAAAATTTTWTATT	170

[illegible][illegible]

Qy	190	TGG	192
		:	!
Db	49	WNG	47

RESULT 5  
AZ523909/c

LOCUS	AZ522909	627 bp	DNA linear	GSS 07-MAY-2000
DEFINITION	213pbB09 Pb MBN #21 Plasmodium berghei genomic 3', genomic survey sequence.			
ACCESSION	AZ522909			

VERSION AZ522909.1 GI:13961380  
KEYWORDS GSS.  
SOURCE plasmodium berghei  
ORGANISM plasmodium berghei

REFERENCE  
AUTHORS  
REMARKS  
1 (bases 1 to 62)  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
Carlton, J.M., R. and Dame, J.B.

**JOURNAL COMMENT**  
Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)  
Contact: Dame JB  
Dept. of Pathobiology, College of Veterinary Medicine

2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA  
Tel: 352 392 4700  
Fax: 352 392 9704

FEATURES

```
1. 0.02
/organism="Plasmodium berghei"
/mol_type="genomic DNA"
/strain="ANKA clone 15cy1 (clone of the ANKA 8417 clone)"
source
```

```

/du_xref="taxon:3621
/dev_stage="asexual blood forms"
/lab_host="Mus musculus"
/clone_lib="pb MEN #21"

```

/note="Vector: pBluescript SK(+)  
excised from lambda ZAP; Site\_1: EcoRV; Site\_2: EcoRV;  
Genomic DNA was prepared from asynchronous blood stage  
forms of the cloned ANKA isolate of *P. berghei* grown in  
vector DNA, pNaegmid

laboratory Swiss white mice. The DNA was purified from contaminating host DNA by Hoechst Dye 33258-CsCl ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38%

formamide at 50°C, as described (Nervick, K.D., Imberski, R.B., and McCutchan, T.P. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size

the EcoRV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform *E. coli* X10-Gold host cells.

BASE COUNT	237 a	92 c	51 g	247 t
ORIGIN				



```
Qy 766 GTGATGCTATCTTTATTATGAGTCGCTCAATCTACACGAGGAGGAGTGTGATTCCAA 825
Db 641 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 826 GAGCTGTCGTGATATCAATGCAATCTTCTGATATACCAACGAGAGAAAGTATTAGTTT 885
Db 581 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 886 ACAACACAGCTAATGGCTACACCACTTAACCTAACGAGTACACCTTACTCAAAAATAAT 945
Db 521 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 946 ACCAAAAGGCTAGACCTCTGCTACTAGGCTAGCTTTTACATCAAAAAGCAATGAC 1005
Db 461 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 1006 TATAGAAAGTAAATACTAGAAAAGCAA 1035
Db 401 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 372

RESULT 8
CNS0187R 878 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN04E04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108993
VERSION AL108993.1 GI:5629297
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
REFERENCE 1 (bases 1 to 878)
AUTHORS Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
TITLE Genoscope.
JOURNAL Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
Location/Qualifiers
1..878
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN04E04"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/notes="end : SP6"

BASE COUNT 279 a 132 c 120 g 182 t 165 others
ORIGIN
Query Match 4.3%; Score 46.8; DB 29; Length 878;
Best Local Similarity 24.6%; Pred. No. 1.7e+02;
Matches 66; Conservative 85; Mismatches 117; Indels 0; Gaps 0;

Qy 781 ATTATGAGTCGCTCCCAATCTACACGAGGAGTGTGATTCGAGAGTGTGCTGAT 840
Db 91 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 150

Qy 841 CAATGCAATCTCTGATAATACCAATCAACGAGAAAGTATTAGTTTACAAACACAGCTAATG 900
Db 151 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 210
```

```
Qy 901 GCTACACCATTAACCTACCGTACACCTACTCAAAAATATACCAAAAGCGTAGAC 960
Db 211 GNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 961 CTCGTGCTCACTAGCGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAGTAAAAA 1020
Db 271 MNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
Qy 1021 TACTGAAAAAGCAATGATTCGCGTCAT 1048
Db 331 TACAAAAAATAAARTWRTAWGAGAAAT 358

RESULT 9
AZ691287/c 881 bp DNA linear GSS 14-DEC-2000
LOCUS ENTMC747R Entamoeba histolytica Sheared DNA Entamoeba histolytica
DEFINITION genomic, genomic survey sequence.
ACCESSION AZ691287
VERSION AZ691287.1 GI:11828433
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
REFERENCE 1 (bases 1 to 881)
AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
TITLE Determination of clone end sequences from Entamoeba histolytica
JOURNAL HM1:IMSS sheared DNA library
COMMENT Unpublished
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
DNA library
Seq primer: M13-Reverse
Class: shotgun
High quality sequence start: 11
High quality sequence stop: 822.

FEATURES
source
Location/Qualifiers
1..881
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt; Site 1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
```

Query Match 4.3%; Score 46.8; DB 28; Length 881;

Best Local Similarity 51.4%; Pred. No. 1.7e+02; Mismatches 108; Conservative 0; Gaps 0;

```
Qy 34 TTTAAATTTTCATATTTTAAAAAACTATTGTAACCTAGTTAAAGTAAGCGTATATAT 93
Db 776 TTTTGGAGTATATTTTCATAATAAAATTTCTATAACTAATAACATAGAGTTATAAA 717
```

94 GGTTAGTTAGCGAAATTAGAAAAGAGGACAAGCATATGAATCTACTTGGATCAAGACGGG 153  
|||  
716 TAATAATAGGTGAGTATTAATAAGTAATCTCATAGAGGAGATTAAATTAATAAGAGAA 657  
|||  
154 TTTTCTTAAAAATGTCGGGTAGTAAAAATTTTCAATCGTAGCTCTTGTATCAGCCAA 213  
|||  
656 GTTCTTATTAAGAAGATTTAATAGGAACATAATAATAGATGTTGTTAAGAAGAGAGAA 597  
|||  
214 TGGCTGTAAACAACAGTCACACTTGAATA 243  
|||  
596 ATAATAAAGAAAGATAGACTGTAATAAGA 567  
|||

RESULT 10	CNS00E0D	1101 bp	DNA	linear	GSS 04-JUN-1999
LOCUS	CNS00E0D				
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR28118 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL067625				
VERSION	AL067625.1	GI:4948725			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of				

COMMENTS

web : [www.genoscope.cns.fr/](http://www.genoscope.cns.fr/)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Oosagawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

FEATURES	FOUND AT	LOCATION/Qualifiers	source
		1. .1101	
		organism="Drosophila melanogaster"	
		mol_type="genomic DNA"	
		db_xref="taxon:7227"	
		clone="BACR28118"	
		clone_lib="RPCI-98"	
		notes="end : T"	
BASE COUNT	360 a	111 c 153 g	300 t
			177 r

Query Match	4.3%	Score 46.6;	DB 29;	Length 1101;
Best Local Similarity	30.2%;	Pred. No. 1.7e+02;		
Matches 88; Conservative	72;	Mismatches 131;	Indels 0;	Gaps 0;

Qy	14	TTTTTCTCCTTACTAATCTCCCTTAAATTTTTCATATTTTAAAAAACTATTGATAAACTA	73
Db	792	TATGDTTAAARMATTMBWTWATRGTTTTATATTTAAAAAAAKGVRAAAWARGR	851
Qy	74	GTTAAGTAAAGCGTATACCTATGTTAGTACGAAATTTAGAAAGAGCAACGACATATGAA	133
Db	852	AWTTTWTWTKKATDIDTAKATTAKGNDDAARRVRAAAVRAGGTACGGTTTWWAWG	911
Qy	134	TCTACTTCGATCAACAGCGGGTTTTTCTAAAAATGTCGGCTAGTAAAAATTTTCAATGGT	193

Db 912 KATASKGTGTATMAAMAGAAHTATTTWTAKWTTAWATATWTRWADAAAAARKTWTATTA 971

Qy 194 AGCTCTTGTAATCAGCCCAATAGCTGTAAACAAGTCACACACTTGAAATAACTGCACATGGC 253

Db 972 YWTTTTTKRKGKGVVVSRVSGSVNSGVRRARVAGAGGDATWAAAASASADAWAA 1031

Qy 254 ACGCAAAACACAGCTCTCAAAATGATGTGTTCTAAATGATGGGGCAAGCAA 304

Db 1032 AAAAABAAATATATTTATTTWAAAATATTTTRKWATEDATWAAADAAAAWAA 1082

RESULT 11

BH132872	BH132872	885 bp	DNA	linear	GSS 07-AUG-2001
LOCUS	ENTNU65STR	Entamoeba histolytica	Sheared DNA	Entamoeba histolytica	
DEFINITION	genomic, genomic survey sequence.				
ACCESSION	BH132872				
VERSION	BH132872.1	GI:15091933			
KEYWORDS	GSS.				
SOURCE	Entamoeba histolytica				
ORGANISM	Entamoeba histolytica				
REFERENCE	Eukaryota; Entamoebidae; Entamoeba.				
AUTHORS	1 (bases 1 to 885)				
TITLE	Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.				
COMMENT	Determination of clone end sequences from Entamoeba histolytica				
JOURNAL	HM1:IMSS sheared DNA library (2001)				
CONTRACT	Unpublished				
CONTACT	Contact: Brendan J Loftus				

## COMMENT

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 3543  
Email: bjloftus@tigr.org  
Clones are derived from the Entamoeba histolytica HM1:IMSS sheared  
DNA library  
Seq primer: M13-Reverse  
Class: shotgun  
High quality sequence start: 18  
High quality sequence stop: 667.

FEATURES  
source

```

/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOS1; Site 1: Bat I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaubin and B.
Barell, Oxford University Press, 1999)."
368 a 96 c 170 g 251 t

```

	Query Match	4.3%	Score 46.4	DB 28	Length 885
	Best Local Similarity	52.0%	Pred. No. 2e+02		
	Matches 104	Conservative 0	Mismatches 96	Indels 0	Gaps 0
QY	44	ATATTTTAAAAAACTATTGATAAATAGTTAAGTAAAGCTATACATGTTAGTAG	103		
Db	6	ATATATTCTAATAAATATTTCTATAAATCTATAAATACATAGATTATAATAATAATAGG	65		
QY	104	CGAAATTTGAAAAGAGCAAGCATATCTACTTGGATCAAGACGGGTTTTTTTCTAA	163		



Db 66 TGAGTATAAATAAGTAATCTCATAGAGGAGATTAAATTAATAAGAGAGATTCCTTATTA 125

Qy 164 AATAATGCGGCTAGTAAATTTTCAATGGTAGCTCTTGTATCAGCCCAATGGCTGAAC 223

Db 126 AAGAATTTAATAGGAACATAAATAATAGATGTTGTTAAGAAGAGAGATAAATAAAG 185

Qy 224 AACAGTCACACTTGAAATA 243

Db 186 AAGATAAGACTGAATAAGA 205

RESULT 12

LOCUS BX344084

DEFINITION BX344084 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI054YA14 3-PRIME, mRNA sequence.

ACCESSION BX344084

VERSION BX344084.1

KEYWORDS GI:30344234

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4553.r For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS1A1014ZC08NP1&cluster=4553.r. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS1A1014ZC08NP1.

FEATURES

source

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSODI054YA14"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 396 a 79 c 211 g 264 t 251 others

ORIGIN

Query Match 4.3%; Score 46.4; DB 13; Length 1201;

Best Local Similarity 30.7%; Pred. No. 1.8e+02;

Matches 101; Conservative 74; Mismatches 154; Indels 0; Gaps 0;

Qy 48 TTTTAAAAAATATTGATAACTAGTTAAGTAAAGCGTATACATATGTTAGTACGAA 107

Db 789 KTTTITADAAWAATTTTKTKATTTTKTKTTTAATAATATTTTTKAATKWTTKAAAGA 848

Qy 108 ATTAGAAAAGGACGACATATGAATCTACTTGGATCAAGCGGTTTTTCTTAAAAA 167

Db 849 AAAKKAAAAAACAACAATAAATAATTTTTTATTTTWTWTAARAATAAAT 908

Qy 168 TGTGCGGTAGTAAAAATTTTCAATGGTAGCTCTTGTATCAGCCCAATGGCTGAACA 227

Db 909 TTTAAATAATTAATAATATWCAAKKKKKADTKTTATTAATAAKDAKCAAGKKKK 968

Qy 228 GTCACTTGAATAATCTGCATGGCAGCAACACACAGCTCTCAATATGTTGTCTA 287

Db 969 KKAATAAAKRAAAAAAARDKDKKDKDAADDKRAKKKKKKDKADKKAADADKDD 1028

Qy 288 AATGATGGCGCAGCAAGTACCTAAACGAGCAATTAGCTTGGACATTCATGACAGTCCT 347

Db 1029 AADKAAAAKKKAKAAAAKKKKKKKKKKKKKKKKKKKKKKKAAAKKKDAKAAKKNA 1088

Qy 348 AACATTATCAAAAACTTTAGTACTAGTCA 376

Db 1089 AAAAANKAAKAAAAADKAAAAAADKKA 1117

RESULT 13

LOCUS AL514423

DEFINITION AL514423 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone CLOBB010ZG02 3-PRIME, mRNA sequence.

ACCESSION AL514423

VERSION AL514423.2

KEYWORDS GI:30464308

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1201)

AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished

COMMENT On Feb 13, 2001 this sequence version replaced gi:12777917.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 10024.f For more information about this cluster, see  
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CLOBB010ZG02FP1&cluster=10024.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CLOBB010ZG02FP1.

FEATURES

Location/Qualifiers

1. .1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CLOBB010ZG02"

/tissue\_type="NEUROBLASTOMA"

/clone\_lib="Homo sapiens NEUROBLASTOMA"

/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

BASE COUNT 361 a 227 c 206 g 259 t 148 others

ORIGIN

Query Match 4.3%; Score 46.2; DB 9; Length 1201;

Best Local Similarity 40.2%; Pred. No. 2e+02;

Matches 84; Conservative 31; Mismatches 94; Indels 0; Gaps 0;

Qy 830 TGTGCTGGTATCAATGCAATCTTCTGATTAATACCATCAACGAGAAAGTATTAGTTTACAA 889

Db 84 TTTTGGTCTWAAAGTAAAMWMTWAAAAAAMWMAAAMAAAAAATTTTTCMAAAAA 143

Qy 890 CACAGCTATGCTACACCATTAATCACTACCATACCGTACACCTACTCAAAAAATAATACCA 949

Db 144 AAAAAAGCAAAWAATGTWAAAAAAGTAAAGAAAAAGAAAAAAGAAAAAAGC 203

Qy 950 AAAGGTAGACCTCTGCTCACTAGGCTAGCTTTTATACATCAAAAAAGCAATGACTATA 1009

Db 204 AAAAAAAGAAWAAAAAATWAAAAAAMAWTCAAAAAAAMAAAAAAMAAAAA 263

Qy 1010 GAAAGTAAAAATCTAGAAAAAGCAATGA 1038



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 5, 2004, 15:08:48 ; Search time 69.8711 Seconds  
(without alignments)  
615.632 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MNLGSRVFSKCRLLVKS.....VNTANGYINVHNGTPTOK 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1414	99.7	271	23	ABP30017 Streptococcus poly
2	1413	99.6	293	16	AA197072 DNA-ase-B. Strept
3	1412	99.6	271	17	AA198823 S. pyogenes DNaseB
4	1393	98.2	271	15	AA198702 Mitogenic factor a
5	1220	86.0	229	16	AA197071 Recombinant DNA-ase
6	456	32.2	91	16	AA197070 DNA-ase-B2 N-termi
7	451	31.8	90	17	AA198821 S. pyogenes DNase
8	203	14.3	43	17	AA198822 Leader peptide. S
9	202	14.2	38	17	AA198825 Recombinant DNaseB

10	182.5	12.9	252	23	ABP27393 Streptococcus poly
11	161.5	11.4	268	23	ABP25596 Streptococcus poly
12	153.5	10.8	263	23	ABP53460 Lactococcus lactis
13	138.5	9.8	354	23	ABP27392 Streptococcus poly
14	138	9.7	285	23	ABP29813 Streptococcus poly
15	134.5	9.5	262	21	AA191138 Group B Streptococ
16	133	9.4	247	20	AA191138 Linoate isomerase
17	133	9.4	247	20	AA191138 L. reuteri linoate
18	131	9.2	157	23	ABP26062 Streptococcus poly
19	131	9.2	242	20	AA192333 Linoate isomerase
20	131	9.2	242	22	ABP6308 L. reuteri linoate
21	131	9.2	261	22	ABP30074 Streptococcus poly
22	130.5	9.2	271	23	ABP27570 Streptococcus poly
23	126.5	8.9	274	21	AA191780 Streptococcus pneu
24	126.5	8.9	274	24	ABU02468 S. pneumoniae type
25	114.5	8.1	1118	25	AA195579 C. histolyticum cl
26	106	7.5	23	17	AA198824 S. pyogenes DNaseB
27	106	7.5	1233	22	AG190558 C. glutamicum prote
28	104.5	7.4	455	19	AA190950 H. pylori ORF 04cp
29	104.5	7.4	455	20	AA191782 H. pylori outer me
30	104.5	7.4	486	18	AA190586 H. pylori cytolas
31	101	7.1	2137	23	ABP39618 Staphylococcus epi
32	100	7.1	160	23	ABP26063 Streptococcus poly
33	98	6.9	522	24	ABP56883 Staphylococcus epi
34	98	6.9	1870	24	ABU19019 Pathogen specific
35	97	6.8	234	22	AAU07562 Trichoderma reesei
36	97	6.8	234	22	AAU07566 Trichoderma reesei
37	97	6.8	234	23	AAU77461 Trichoderma reesei
38	96	6.8	234	23	AAU77456 Trichoderma reesei
39	95	6.7	234	23	AAU77452 Trichoderma reesei
40	95	6.7	234	23	AAU77457 Trichoderma reesei
41	94	6.6	234	23	AAU77455 Trichoderma reesei
42	94	6.6	234	23	AAU77465 Trichoderma reesei
43	94	6.6	293	20	AA191365 Wild-type Staphylo
44	94	6.6	595	21	AA192370 Escherichia coli F
45	94	6.6	621	21	AA1959201 Bacillus sp. AA386

## ALIGNMENTS

RESULT 1  
ABP30017  
ID ABP30017 standard; Protein; 271 AA.

AC ABP30017;

DT 02-JUL-2002 (first entry)

XX Streptococcus polypeptide SEQ ID NO 9210.

DE Streptococcus; GAS; CBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antinflammatory; infection; vaccine; meningitis; gene therapy.

XX Streptococcus pyogenes.

PN WO200234771-A2.

PD 02-MAY-2002.

XX 29-OCT-2001; 2001WO-GB04789.

PR 27-OCT-2000; 2000GB-0026333.

PR 24-NOV-2000; 2000GB-0028727.

PR 07-MAR-2001; 2001GB-0005640.

XX (CHIR-) CHIRON SPA.

PA (GENO-) INST GENOMIC RES.

XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelin H;

XX

DR WPI; 2002-352536/38.  
DR N-PSDB; ABN70648.  
XX  
PT New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PT for detecting a compound that binds to the protein -  
XX  
PS Claim 1; Page 4045; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
SQ Sequence 271 AA;  
Query Match 99.7%; Score 1414; DB 23; Length 271;  
Best Local Similarity 99.6%; Pred. No. 3.5e-138;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVNDGASKY 60  
Db 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVNDGASKY 60  
Qy 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARTGLTYANVEG 120  
Db 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARTGLTYANVEG 120  
Qy 121 SYGVRQSGKQKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180  
Db 121 SYGVRQSGKQKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180  
Qy 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRGYLYYEVAPIYNADELIPRAVTVSMQ 240  
Db 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRGYLYYEVAPIYNADELIPRAVTVSMQ 240  
Qy 241 SSDNTINEKVLVYNTANGYTYNYHNGTPTQK 271  
Db 241 SSDNTINEKVLVYNTANGYTYNYHNGTPTQK 271  
RESULT 2  
AAR70702  
ID AAR70702 standard; Protein; 293 AA.  
XX AC AAR70702;  
XX  
DT 25-MAR-2003 (updated)  
DT 15-AUG-1995 (first entry)  
XX  
DE DNA-ase-B.  
XX  
XX DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.  
XX  
XX Streptococcus pyogenes (ATCC 14289).  
XX  
FH Key Location/Qualifiers  
FT Protein 46...293  
FT Peptide /note= "mature protein"  
FT 1..45

FT misc\_difference 274 /note= "leader peptide: claim 11"  
FT FT /note= "in-frame stop codon"  
XX  
PN W09500650-A1.  
XX  
PD 05-JAN-1995.  
XX  
XX 18-MAY-1994; 94WO-US05626.  
XX  
PR 23-JUN-1993; 93US-0082845.  
XX  
XX (BECI ) BECKMAN INSTR INC.  
PA  
XX Adams CW, Belei CM, Pang PPy;  
PI WPI; 1995-052087/07.  
XX N-PSDB; AAQ85037.  
DR  
XX New DNA encoding Streptococcus pyogenes DNase B - for diagnosing  
PT S. pyogenes infection, also new promoter for expressing other  
PT proteins  
XX  
PS Disclosure; Fig 5; 97pp; English.  
XX  
XX The sequence shows a gene product corresponding to a Streptococcus  
CC pyogenes DNA-ase-B. The protein is useful as a diagnostic agent,  
CC vaccine or as an aerosol to treat excessive lung viscosity, e.g. in  
CC cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 293 AA;  
Query Match 99.6%; Score 1413; DB 16; Length 293;  
Best Local Similarity 100.0%; Pred. No. 5e-138;  
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 NLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVNDGASKYL 61  
Db 4 NLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVNDGASKYL 63  
Qy 62 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARTGLTYANVEG 121  
Db 64 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARTGLTYANVEG 123  
Qy 122 YGVRQSGKQKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNAV 181  
Db 124 YGVRQSGKQKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNAV 183  
Qy 182 TGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRGYLYYEVAPIYNADELIPRAVTVSMQ 241  
Db 184 TGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRGYLYYEVAPIYNADELIPRAVTVSMQ 243  
Qy 242 SDNTINEKVLVYNTANGYTYNYHNGTPTQK 271  
Db 244 SDNTINEKVLVYNTANGYTYNYHNGTPTQK 273  
RESULT 3  
AAR88823  
ID AAR88823 standard; Protein; 271 AA.  
XX AC AAR88823;  
XX  
XX 25-JUN-1996 (first entry)  
DT  
XX S. pyogenes DNaseB and leader sequence.  
DE  
XX DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
XX diagnosis; Escherichia coli.  
XX  
XX Streptococcus pyogenes strain ATCC 14289.  
XX

PH Key Location/Qualifiers  
FT Peptide 1..43  
FT Protein /label= Sig\_peptide  
FT 44..271  
PN /label= Mat\_protein  
XX WO9606174-A1.  
XX 29-FEB-1996.  
XX 18-AUG-1994; 94WO-US09450.  
XX 18-AUG-1994; 94WO-US09450.  
XX (BECI) BECKMAN INSTR INC.  
XX Adams CW, Belei MC, Pang PPV;  
XX WPI; 1996-151377/15.  
XX N-PSDB; AAT12774.  
XX New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
XX prodn. of the enzyme in other bacteria, useful in immunoassays or  
XX for treating cystic fibrosis  
XX Claim 1; Page 67-70; 115pp; English.  
XX Streptococcus pyogenes DNase B, including the leader peptide,  
XX has the amino acid sequence given in AAR88921. The enzyme can  
XX be obd. on a large scale by expression of encoding DNA (AAT12774)  
XX in transformed host cells, esp. Escherichia coli. Inclusion  
XX of the leader peptide facilitates purification of the recombinant  
XX enzyme. The DNase B is useful in immunoassays to detect  
XX anti-DNase B antibodies in serum as a marker for S. pyogenes  
XX infection, and is also useful as a vaccine or for treatment, via  
XX aerosol delivery, of cystic fibrosis.  
XX SQ Sequence 271 AA;  
Query Match 99.6%; Score 1412; DB 17; Length 271;  
Best Local Similarity 99.6%; Pred. No. 5.6e-138;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALARQTQVSNVDVNLNDGASKY 60  
Db 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALARQTQVSNVDVNLNDGASKY 60  
Qy 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120  
Db 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120  
Qy 121 SYGVRQSGFNQNPAGWTGNPNHVYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNA 180  
Db 121 SYGVRQSGFNQNPAGWTGNPNHVYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNA 180  
Qy 181 VTGTRTQNVGRDQKGMRYTEQRAQEWLEANDRGYLYVEVAPTYNADELIIPRAVVVSMQ 240  
Db 181 VTGTRTQNVGRDQKGMRYTEQRAQEWLEANDRGYLYVEVAPTYNADELIIPRAVVVSMQ 240  
Qy 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271  
Db 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271  
RESULT 4  
AAR58702  
ID AAR58702 standard; Protein; 271 AA.  
XX AAR58702;  
XX AAR58702;  
DT 25-MAR-2003 (updated)  
DT 29-MAR-1995 (first entry)  
XX

DE Mitogenic factor associated with group A Streptococci.  
XX mitogenic factor; microdetection; group A streptococci; spe;  
XX erythrogenic toxin; streptococcal pyrogenic exotoxin; blastogens;  
KW scarlet fever toxin; erythematous skin reaction; infectious disease;  
KW delayed hypersensitivity; ss.  
XX Streptococcus pyogenes.  
OS EP613947-A2.  
XX 07-SEP-1994.  
PD 31-JAN-1994; 94EP-0101386.  
XX 01-FEB-1993; 93JP-0037383.  
PR (SHIO) SHIONOGI & CO LTD.  
PA Hara A, Hinuma Y, Igarashi H, Iwasaki M, Kishishita M;  
XX Okumura K, Takeda Y, Yutsudo T;  
PI WPI; 1994-272994/34.  
XX N-PSDB; AAQ71612.  
DR New mitogenic factor gene from Streptococcus pyogenes - used to  
XX develop prods. for the early diagnosis of infectious disease  
XX caused by gp A streptococci  
XX Claim 7; Page 12-13; 20pp; English.  
XX AAR58702 shows a mitogenic factor which exhibits rabbit peripheral  
XX blood lymphocyte mitogenicity and/or DNA hydrolysing activity. It is  
XX strongly associated with group A Streptococci and the nucleotide  
XX sequences can be used for the microdetection of the gene and provide  
XX an early diagnosis of infectious disease caused by the bacteria.  
XX (See also AAQ71613-26).  
XX (Updated on 25-MAR-2003 to correct PN field.)  
XX SQ Sequence 271 AA;  
Query Match 98.2%; Score 1393; DB 15; Length 271;  
Best Local Similarity 98.5%; Pred. No. 5.3e-136;  
Matches 267; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALARQTQVSNVDVNLNDGASKY 60  
Db 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALARQTQVSNVDVNLNDGASKY 60  
Qy 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120  
Db 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120  
Qy 121 SYGVRQSGFNQNPAGWTGNPNHVYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNA 180  
Db 121 SYGVRQSGFNQNPAGWTGNPNHVYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNA 180  
Qy 181 VTGTRTQNVGRDQKGMRYTEQRAQEWLEANDRGYLYVEVAPTYNADELIIPRAVVVSMQ 240  
Db 181 VTGTRTQNVGRDQKGMRYTEQRAQEWLEANDRGYLYVEVAPTYNADELIIPRAVVVSMQ 240  
Qy 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271  
Db 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271  
RESULT 5  
AAR70701  
ID AAR70701 standard; protein; 229 AA.  
XX AAR70701;  
XX AAR70701;  
DT 25-MAR-2003 (updated)  
XX

DT 14-AUG-1995 (first entry)  
XX  
DE Recombinant DNA-ase-B.  
XX  
KW DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.  
XX  
OS Streptococcus pyogenes.  
XX  
PN WO9500650-A1.  
XX  
PD 05-JAN-1995.  
XX  
PF 18-MAY-1994; 94WO-US05626.  
XX  
PR 23-JUN-1993; 93US-0082845.  
XX  
PA (BECI ) BECKMAN INSTR INC.  
XX  
PI Adams CW, Belei CM, Pang PPY;  
XX  
DR WPI; 1995-052087/07.  
XX  
PT New DNA encoding Streptococcus pyogenes Dnaase B - for diagnosing  
PT S. pyogenes infection, also new promoter for expressing other  
PT proteins  
XX  
PS Claim 1; Fig 4; 97pp; English.  
XX  
CC This is the full-length sequence of Streptococcus pyogenes  
CC DNA-ase-B, which is a marker of S. pyogenes infection. The protein  
CC may be used as a diagnostic agent or vaccine for S. pyogenes, or  
CC may be used as an aerosol to treat excessive lung viscosity, e.g.  
CC in cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 229 AA;  
  
Query Match 86.0%; Score 1220; DB 16; Length 229;  
Best Local Similarity 100.0%; Pred. No. 3.8e-118;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 43 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALPKAGDILYKLDE 102  
Db 1 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALPKAGDILYKLDE 60  
  
QY 103 LGRTTARGTLTYANVEGSGVROSGKQNPAGTGNPNHVKYKIEWNLGLSVYGFVN 162  
Db 61 LGRTTARGTLTYANVEGSGVROSGKQNPAGTGNPNHVKYKIEWNLGLSVYGFVN 120  
  
QY 163 RSHLIADSLGDLALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEANRDGVLYYEVA 222  
Db 121 RSHLIADSLGDLALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEANRDGVLYYEVA 180  
  
QY 223 PIYNADLIPRAVNVVWQSSDNTINEKVLVYNTANGTYINHGTPQK 271  
Db 181 PIYNADLIPRAVNVVWQSSDNTINEKVLVYNTANGTYINHGTPQK 229  
  
RESULT 6  
AAR70700  
ID AAR70700 standard; Protein; 91 AA.  
XX  
AC AAR70700;  
XX  
XX 25-MAR-2003 (updated)  
DT 14-AUG-1995 (first entry)  
XX  
XX DNA-ase-B2 N-terminal fragment.  
DE DNA-ase-B2 N-terminal fragment; diagnostic; vaccine; cystic  
XX fibrosis therapy.  
XX  
OS Streptococcus pyogenes.

XX WO9500650-A1.  
PN 05-JAN-1995.  
PD 18-MAY-1994; 94WO-US05626.  
XX  
PF 23-JUN-1993; 93US-0082845.  
XX  
PR (BECI ) BECKMAN INSTR INC.  
XX  
PA Adams CW, Belei CM, Pang PPY;  
XX  
PI WPI; 1995-052087/07.  
XX  
DR N-PSDB; AAQ85036.  
XX  
PT New DNA encoding Streptococcus pyogenes Dnaase B - for diagnosing  
PT S. pyogenes infection, also new promoter for expressing other  
PT proteins  
XX  
PS Claim 3; Fig 3; 97pp; English.  
XX  
CC The sequence represents the N-terminal fragment of Streptococcus  
CC pyogenes DNA-ase-B, which is a marker of S. pyogenes infection.  
CC The protein may be used as a diagnostic agent or vaccine for S.  
CC pyogenes, or may be used as an aerosol to treat excessive lung  
CC viscosity, e.g. in cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 91 AA;  
  
Query Match 32.2%; Score 456; DB 16; Length 91;  
Best Local Similarity 100.0%; Pred. No. 2.2e-39;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MNLGSRVPSKCRVLKFSWALVSATMAVTTVTLENTALARQTQVNDVLDGASKY 60  
Db 1 MNLGSRVPSKCRVLKFSWALVSATMAVTTVTLENTALARQTQVNDVLDGASKY 60  
  
QY 61 LNEALAWTFNDSPNYKTLGTSQITPALPK 91  
Db 61 LNEALAWTFNDSPNYKTLGTSQITPALPK 91  
  
RESULT 7  
AAR88821  
ID AAR88821 standard; Protein; 90 AA.  
XX  
AC AAR88821;  
XX  
DT 25-JUN-1996 (first entry)  
XX  
XX S. pyogenes Dnaase B partial sequence.  
DE Dnaase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
KW diagnosis; Escherichia coli.  
XX  
OS Streptococcus pyogenes strain ATCC 14289.  
XX  
XX  
FH Key Location/Qualifiers  
FT Peptide 1.43  
FT /label= sig\_peptide  
XX  
PN WO9606174-A1.  
XX  
PD 29-FEB-1996.  
XX  
XX 18-AUG-1994; 94WO-US09450.  
PF 18-AUG-1994; 94WO-US09450.  
PR 18-AUG-1994; 94WO-US09450.  
XX  
PA (BECI ) BECKMAN INSTR INC.  
XX

PI Adams CW, Belel MC, Pang PPY;  
 XX WPI; 1996-151377/15.  
 DR N-PSDB; AAT12773.  
 XX  
 PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
 PT prodn. of the enzyme in other bacteria, useful in immunoassays or  
 PT for treating cystic fibrosis  
 XX  
 PS Example 2; Fig 3; 115pp; English.  
 XX  
 CC A partial sequence (AAR88821) for Streptococcus pyogenes DNase B  
 CC is the product of DNA clone lambda 2-6 (AAT12773) isolated from a  
 CC S. pyogenes DNA library. A full-length sequence is given in  
 CC AAR88823. The insert of the DNase B-encoding DNA clone can be  
 CC transferred to expression plasmids for large-scale recombinant  
 CC DNase prodn. in E. coli or other bacterial hosts. It is used in  
 CC immunoassays to detect anti-DNase B antibodies in serum as a  
 CC marker of S. pyogenes infection, and is also useful as a vaccine  
 CC or for cystic fibrosis treatment via aerosol delivery.  
 XX  
 SQ Sequence 90 AA;  
 Query Match 31.8%; Score 451; DB 17; Length 90;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-39;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60  
 DB 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60  
 QY 61 LNEALAWTFNDSPNYKTLGTSGITPALFP 90  
 DB 61 LNEALAWTFNDSPNYKTLGTSGITPALFP 90  
 RESULT 8  
 AAR88822  
 ID AAR88822 standard; Peptide; 43 AA.  
 XX  
 AC AAR88822;  
 XX  
 DT 25-JUN-1996 (first entry)  
 XX  
 DE Leader peptide.  
 XX  
 KW DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
 KW diagnosis; leader peptide.  
 XX  
 OS Streptococcus pyogenes strain ATCC 14289.  
 XX  
 PN WO9606174-A1.  
 XX  
 PD 29-FEB-1996.  
 XX  
 PF 18-AUG-1994; 94WO-US09450.  
 XX  
 PR 18-AUG-1994; 94WO-US09450.  
 XX  
 PA (BECI ) BECKMAN INSTR INC.  
 XX  
 PI Adams CW, Belel MC, Pang PPY;  
 XX  
 DE WPI; 1996-151377/15.  
 XX  
 PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
 PT prodn. of the enzyme in other bacteria, useful in immunoassays or  
 PT for treating cystic fibrosis  
 XX  
 PS Claim 15; Page 89; 115pp; English.  
 XX  
 CC A leader peptide (AAR88822) of Streptococcus pyogenes DNase B  
 CC can be used to facilitate prodn. of recombinant DNase B (see

CC AAR88823) in transformed bacterial hosts, e.g. Escherichia coli.  
 CC The leader peptide can also be used for expression and prodn.  
 CC of other recombinant proteins in bacteria. The product is  
 CC excreted by the host into the culture medium and is easily  
 CC recovered.  
 XX  
 SQ Sequence 43 AA;  
 Query Match 14.3%; Score 203; DB 17; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-13;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 43  
 DB 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 43  
 RESULT 9  
 AAR88825  
 ID AAR88825 standard; Peptide; 38 AA.  
 XX  
 AC AAR88825;  
 XX  
 DT 25-JUN-1996 (first entry)  
 XX  
 DE Recombinant DNaseB N-terminal peptide.  
 XX  
 KW DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
 KW diagnosis; Streptococcus pyogenes.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9606174-A1.  
 XX  
 PD 29-FEB-1996.  
 XX  
 PF 18-AUG-1994; 94WO-US09450.  
 XX  
 PR 18-AUG-1994; 94WO-US09450.  
 XX  
 PA (BECI ) BECKMAN INSTR INC.  
 XX  
 PI Adams CW, Belel MC, Pang PPY;  
 XX  
 DE WPI; 1996-151377/15.  
 XX  
 PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
 PT prodn. of the enzyme in other bacteria, useful in immunoassays or  
 PT for treating cystic fibrosis  
 XX  
 PS Example 6; Page 48; 115pp; English.  
 XX  
 CC The N-terminal sequence (AAR88825) of a Streptococcus pyogenes  
 CC recombinant DNase B expressed in Escherichia coli transformants was  
 CC detd. A modified DNase B gene (AAT12780) was produced by PCR  
 CC amplification of the native DNase B gene (AAT12774) in clone lambda 2-6.  
 CC The gene was inserted into vector del-33 for expression in E. coli.  
 CC Modification of the gene resulted in the presence of an additional  
 CC Arg residue at the DNase B N-terminus (see also AAR88824).  
 XX  
 SQ Sequence 38 AA;  
 Query Match 14.2%; Score 202; DB 17; Length 38;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-13;  
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 43 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
 DB 1 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 38  
 RESULT 10  
 ABP27393





Best Local Similarity 25.6%; Pred. No. 4.7e-08;  
Matches 60; Conservative 36; Mismatches 83; Indels 55; Gaps 12;

QY 21 MVALVSATMAVTTVTLENTALARQTVSDVNDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
D 15 VVILIAILTTTSTVTVAARKIRNFPDPTTEILGKATE----- 53  
QY 81 TSQITPALFPKAG--DILYSKLDLGRTRTARGTLTYANV-----EGSYGVRSQSPGKNQNP 134  
D 54 ----TPGILPFTGSYQLVGLDNLQRP-----TFAHIQKQDQDFBNIKRK-GLKFPN 101  
QY 135 AGMTGNPNHVKYKLEWNLGLSYVGDFW--NRSHLIADSLGG--DALRVNAVGTGRTONVG 190  
D 102 PGW-----H-NYKLT-DANGKT-----TWLMDRGHLVGVQFSGLNDEPK-NLVTMTKYLNTG 150  
QY 191 GRDQKG-GMRYTEQRAQEWLEBANRDGYLYEVAPIYNADELIPRAVVVSMQSSD 243  
D 151 FSDKNPLGLYENRLDLSWLAHFNWLDYKVTVPYHKNELVPRQVQLVYVGID 204

RESULT 12  
ABB53460  
ID ABB53460 standard; Protein; 263 AA.  
XX AC ABB53460;  
XX DT 16-MAY-2002 (first entry)  
XX DE Lactococcus lactis protein ybfB.  
XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.  
XX OS Lactococcus lactis IL1403.  
XX PN FR2807446-A1.  
XX PD 12-OCT-2001.  
XX PF 11-APR-2000; 2000FR-0004630.  
XX PR 11-APR-2000; 2000FR-0004630.  
XX PA (INRG ) INRA INST NAT RECH AGRONOMIQUE.  
XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
XX PS WPI; 2002-043418/06.  
XX PT New nucleotide sequence useful in the identification or Lactococcus  
XX lactic acid and related species -  
XX PS Claim 6; SEQ ID No 162; 2504pp; French.

CC The present invention is related to a Lactococcus lactis nucleotide  
CC sequence (AB90521) and related proteins (ABB53300-ABB55621). The  
CC nucleic acid sequence is useful in the detection and/or amplification of  
CC nucleic acid sequence, particularly to identify Lactococcus lactis or  
CC related species. The proteins of the invention are useful for the  
CC biosynthesis or biodegradation of a composition of interest. The  
CC invention helps research in lactic bacteria, particularly useful in the  
CC production of yogurt and cheese.  
CC Note: The sequence data for this patent is based on equivalent patent  
CC WO200177334 (published 18-OCT-2001) which is available in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 263 AA;

Query Match 10.8%; Score 153.5; DB 23; Length 263;  
Best Local Similarity 24.4%; Pred. No. 3.1e-07;  
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;

QY 68 TFNDSPNYKTLGTS-QITPALFPKAGDILYSKLDLGR-----RTARGTL 113  
D 151 FSDKNPLGLYENRLDLSWLAHFNWLDYKVTVPYHKNELVPRQVQLVYVGID 204

Db 39 TDNSSQVSTKSLASSVKQAPLTFNQRMQWANTDALGRAVDSHIQLKDSQEPKVKREPL 98  
QY 114 TYANVEGSGVRSQFGKNQNPAGTGNPNHVKYKLEWNLGLSYVGDFW--NRSHLIA--- 168  
D 99 TY-----H-NYFYFKSDSGIGKMWLMARGHLVGVQF 135  
QY 169 DSLGGDALRV-----NAVGTGRTONVGGRDQKGMRYTEQRAQEWLEBANRDGYLY 219  
D 136 SGLNNEARNLVPETAMFNGNFTGTNDGNT-----ASMLYENRLDLSWLANHPNYLYDY 189  
QY 220 EVAPIYNADLIPRAV--VVSMSQSDNTINEK-----VLVYNTANGYTI 261  
D 190 QVTPLVEGNELLPQRLAYVGDKNQTLISIKLGGREKSGNGGATVVVDNVAFNAKI 249  
QY 262 NYHNGT 267  
D 250 NYADGT 255

RESULT 13  
ABP27392  
ID ABP27392 standard; Protein; 354 AA.  
XX AC ABP27392;  
XX DT 02-JUL-2002 (first entry)  
XX DE Streptococcus polypeptide SEQ ID NO 3960.  
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
XX group A streptococcus; Streptococcus pyogenes; antibacterial;  
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX OS Streptococcus agalactiae.  
XX PN WO200234771-A2.  
XX PD 02-MAY-2002.  
XX PF 29-OCT-2001; 2001WO-GB04789.  
XX PR 27-OCT-2000; 2000GB-0026333.  
XX PR 24-NOV-2000; 2000GB-0028727.  
XX PR 07-MAR-2001; 2001GB-0005640.  
XX PA (CHIR-) CHIRON SPA.  
XX PA (GENO-) INST GENOMIC RES.  
XX PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;  
XX PI Tettelin H;  
XX PS WPI; 2002-352536/38.  
XX DR N-PSDB; ABN68023.  
XX PT New Streptococcus protein for the treatment or prevention of infection  
XX or disease caused by Streptococcus bacteria, such as meningitis, and  
XX for detecting a compound that binds to the protein -  
XX PS Claim 1; Page 3550; 4525pp; English.

CC The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (1), nucleic acids encoding (1), ABN68044-ABN71326 and  
CC antibodies that bind (1) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (1) are used to detect Streptococcus in a  
CC biological sample. (2) is used to determine whether a compound binds to  
CC (1). A composition comprising (1) or a nucleic acid encoding (1), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic

[illegible]

XX (MICR-) MICROBIAL TECHNIKS LTD.  
XX Le Page RWF, Wells JM, Hanniffy SB;  
XX WPI; 2000-195299/17.  
XX  
XX New Group B Streptococcus protein, useful as vaccine, for diagnosis of  
PT Streptococcal infections and for screening of antibodies or affibodies  
XX  
XX Claim 1; Fig 1; 123pp; English.  
XX  
XX AAA05803 to AAA05872 encode proteins, polypeptides and peptides (given  
CC in AAY91275 to AAY91343) isolated from Group B Streptococcus (GBS), also  
CC known as Streptococcus agalactiae. The GBS polynucleotides and  
CC polypeptides have antibacterial activity. Immunogenic compositions  
CC comprising GBS polynucleotides or polypeptides can be used as vaccines  
CC and for the treatment or prophylaxis of GBS infection. The  
CC polynucleotides and polypeptides can also be used in the detection of GBS  
CC and for screening DNA encoding bacterial cell envelope associated or  
CC secreted antigens in gram positive bacteria. AAA05873 to AAA05941  
CC represent primers used in the exemplification of the present invention.  
XX  
SQ Sequence 262 AA;

Query Match 9.5%; Score 134.5; DB 21; Length 262;  
Best Local Similarity 25.8%; Pred. No. 2.9e-05;  
Matches 39; Conservative 25; Mismatches 60; Indels 27; Gaps 5;  
Qy 131 NONPAGW--TCGNPHVKYKIEWLNGLSYGVDFWNRSHLIADSLGGDALR-----VN 179  
Db 125 NWKPLGWHQVATNDHYGHVD-----KGLIAYALAGNNFKGWDASVSNPQN 171  
Qy 180 AVTGTRTQNVGGRDQKGMYTEQRAQEWLEARNRDGYLYYEVAPIYNAD-ELIPRAVVVS 238  
Db 172 VVTQTAHSNQSQKINRGQNYYESLVKAVDQNK--RVRVRYTPLYRNDTDLVFFAMHLE 229  
Qy 239 MQSSDNTINEKVLVYNTANGYTINYHNGTPT 269  
Db 230 AKSQDGTLEFPNVAIPNTQASVTMDYATGEIT 260

Search completed: January 5, 2004, 18:38:08  
Job time : 72.8711 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:35:50 ; Search time 30.9429 Seconds  
(without alignments)  
370.561 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MNLGSRVFSKCKRLVKFS.....VNTANGYTYNHVNGTPTOK 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	271	4 US-08-393-889-8	Sequence 8, Appli
2	1418	100.0	271	5 PCT-US94-09450-8	Sequence 8, Appli
3	1397.5	98.6	272	4 US-08-393-889-15	Sequence 15, Appli
4	1397.5	98.6	272	5 PCT-US94-09450-15	Sequence 15, Appli
5	1220	86.0	229	4 US-08-393-889-9	Sequence 9, Appli
6	1220	86.0	229	5 PCT-US94-09450-9	Sequence 9, Appli
7	203	14.3	43	4 US-08-393-889-1	Sequence 1, Appli
8	203	14.3	43	5 PCT-US94-09450-1	Sequence 1, Appli
9	202	14.2	38	4 US-08-393-889-6	Sequence 6, Appli
10	202	14.2	38	5 PCT-US94-09450-6	Sequence 6, Appli
11	170	12.0	32	4 US-08-393-889-16	Sequence 16, Appli
12	170	12.0	32	5 PCT-US94-09450-16	Sequence 16, Appli
13	114.5	8.1	1118	3 US-09-379-523-3	Sequence 3, Appli
14	113	8.0	97	4 US-09-107-532A-5038	Sequence 5038, Ap
15	106	7.5	23	4 US-08-393-889-4	Sequence 4, Appli
16	106	7.5	23	5 PCT-US94-09450-4	Sequence 4, Appli
17	101	7.1	2137	4 US-09-134-001C-4463	Sequence 4463, Ap
18	99.5	7.0	858	4 US-09-252-991A-29756	Sequence 29756, A
19	95.5	6.7	379	4 US-09-107-532A-6904	Sequence 6904, Ap
20	94	6.6	621	4 US-09-311-626B-2	Sequence 2, Appli
21	92	6.5	234	1 US-08-032-848C-9	Sequence 9, Appli
22	92	6.5	234	1 US-08-438-870-9	Sequence 9, Appli
23	92	6.5	234	3 US-09-146-770-3	Sequence 3, Appli
24	92	6.5	234	3 US-09-216-295-3	Sequence 3, Appli
25	92	6.5	234	4 US-09-633-084-3	Sequence 3, Appli
26	92	6.5	234	4 US-10-075-872-3	Sequence 3, Appli
27	92	6.5	234	4 US-10-261-997-3	Sequence 3, Appli

Sequence 15, Appli  
Sequence 2, Appli  
Sequence 30579, A  
Sequence 46, Appli  
Sequence 9, Appli  
Sequence 4570, Ap  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 49, Appli  
Sequence 47, Appli  
Sequence 45, Appli  
Sequence 43, Appli  
Sequence 5276, Ap  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 24, Appli  
Sequence 24, Appli

ALIGNMENTS

RESULT 1  
US-08-393-889-8  
; Sequence 8, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-393-889-8

Query Match 100.0%; Score 1418; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.3e-142;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNLGSRVFSKCKRLVKFSMVALVSATWAVTTVTLENTALARTQVSNVDVNDGASKY 60

Db 1 MNLGSRVFSKCRVLFVSMVALVSATMAVTTVLTALAQVQVNDVVLNDGASKY 60  
QY 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 120  
Db 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 120  
QY 121 SYGVRSQFGKQNPAGWTGNPNHVYKIEWLNGLSYVGD FWNRSHLIADSLGGDALRVNA 180  
Db 121 SYGVRSQFGKQNPAGWTGNPNHVYKIEWLNGLSYVGD FWNRSHLIADSLGGDALRVNA 180  
QY 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLYYEYVAPIYNADELIPRAVVVSMQ 240  
Db 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLYYEYVAPIYNADELIPRAVVVSMQ 240  
QY 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTOK 271  
Db 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTOK 271

RESULT 2  
PCT-US94-09450-8  
; Sequence 8, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beckman Instruments, Inc.  
; STREET: 2500 Harbor Boulevard  
; CITY: Fullerton  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92634  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09450  
; FILING DATE: 18-AUG-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/082,845  
; FILING DATE: 23-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: May, William H.  
; REGISTRATION NUMBER: 26,769  
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-09450-8

Query Match 100.0%; Score 1418; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 1.3e-142;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNLGSRVFSKCRVLFVSMVALVSATMAVTTVLTALAQVQVNDVVLNDGASKY 60  
Db 1 MNLGSRVFSKCRVLFVSMVALVSATMAVTTVLTALAQVQVNDVVLNDGASKY 60  
QY 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 120

Db 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 120  
QY 121 SYGVRSQFGKQNPAGWTGNPNHVYKIEWLNGLSYVGD FWNRSHLIADSLGGDALRVNA 180  
Db 121 SYGVRSQFGKQNPAGWTGNPNHVYKIEWLNGLSYVGD FWNRSHLIADSLGGDALRVNA 180  
QY 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLYYEYVAPIYNADELIPRAVVVSMQ 240  
Db 181 VTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLYYEYVAPIYNADELIPRAVVVSMQ 240  
QY 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTOK 271  
Db 241 SSDNTINEKVLVYNTANGTYTINYHNGTPTOK 271

RESULT 3  
US-08-393-889-15  
; Sequence 15, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-393-889-15

Query Match 98.6%; Score 1397.5; DB 4; Length 272;  
Best Local Similarity 99.6%; Pred. No. 2e-140;  
Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2 NLGSRVFSKCRVLFVSMVALVSATMAVTTVLTALAQVQVNDVVLNDGASKYL 61  
Db 4 NLGSRVFSKCRVLFVSMVALVSATMAVTTVLTALAQVQVNDVVLNDGASKYL 62  
QY 62 NEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 121  
Db 63 NEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEG 122

QY 122 YGVRQSGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWNRLHLSIADSLGGDALRVNAV 181  
DB 123 YGVRQSGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWNRLHLSIADSLGGDALRVNAV 182  
QY 182 TGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVAPIYNADELIPRAVVVSMQS 241  
DB 183 TGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVAPIYNADELIPRAVVVSMQS 242  
QY 242 SDNTINEKVLVYNTANGTYTINYHNGTPTOK 271  
DB 243 SDNTINEKVLVYNTANGTYTINYHNGTPTOK 272

RESULT 4  
PCT-US94-09450-15  
; Sequence 15, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beckman Instruments, Inc.  
; STREET: 2500 Harbor Boulevard  
; CITY: Fullerton  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92634  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09450  
; FILING DATE: 18-AUG-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/082,845  
; FILING DATE: 23-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: May, William H.  
; REGISTRATION NUMBER: 26,769  
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 796-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US94-09450-15

Query Match 98.6%; Score 1397.5; DB 5; Length 272;  
Best Local Similarity 99.8%; Pred. No. 2e-140; Mismatches 0; Indels 1; Gaps 1;  
Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 NLGSRVFSKCRVLFKFSVALVSATMAVTTVTLENTALARTQVSDVNDVLDGASKYL 61  
DB 4 NLGSRVFSKCRVLFKFSVALVSATMAVTTVTLENTALA-QTVSDVNDVLDGASKYL 62  
QY 62 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDDELGRTRTARTGLTYANVEGS 121  
DB 63 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDDELGRTRTARTGLTYANVEGS 122  
QY 122 YGVRQSGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWNRLHLSIADSLGGDALRVNAV 181  
DB 123 YGVRQSGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWNRLHLSIADSLGGDALRVNAV 182

QY 182 TGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVAPIYNADELIPRAVVVSMQS 241  
DB 183 TGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVAPIYNADELIPRAVVVSMQS 242  
QY 242 SDNTINEKVLVYNTANGTYTINYHNGTPTOK 271  
DB 243 SDNTINEKVLVYNTANGTYTINYHNGTPTOK 272

RESULT 5  
US-08-393-889-9  
; Sequence 9, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; US-08-393-889-9

Query Match 86.0%; Score 1220; DB 4; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.2e-121; Mismatches 0; Indels 0; Gaps 0;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ROTQVSDNVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLD 102  
DB 1 ROTQVSDNVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLD 60  
QY 103 LGRTRTARTGLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWN 162  
DB 61 LGRTRTARTGLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKYKLEWNLGLSYVGD FWN 120  
QY 163 RSHLTADSLGGDALRVNAVTTGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVA 222  
DB 121 RSHLTADSLGGDALRVNAVTTGTRTQNVGGRDQGGMRVTEQRAQEWLEARNRQDGLYYEVA 180

QY 223 PYNADLIPRAVVMQSSDNTINEKLVYNTANGYTYNHGTPQK 271  
Db 181 PYNADLIPRAVVMQSSDNTINEKLVYNTANGYTYNHGTPQK 229

RESULT 6  
PCT-US94-09450-9  
; Sequence 9, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beckman Instruments, Inc.  
; STREET: 2500 Harbor Boulevard  
; CITY: Fullerton  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92634  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09450  
; FILING DATE: 18-AUG-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/082,845  
; FILING DATE: 23-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: May, William H.  
; REGISTRATION NUMBER: 26,769  
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
PCT-US94-09450-9

Query Match 86.0%; Score 1220; DB 5; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.2e-121; Indels 0; Gaps 0;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLE 102  
Db 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLE 60  
QY 103 LGRTRTARGTLTYANVGSYGVRSFGKQNPAGTGNPHVYKIKIWLNLGLSYVGDFFN 162  
Db 61 LGRTRTARGTLTYANVGSYGVRSFGKQNPAGTGNPHVYKIKIWLNLGLSYVGDFFN 120  
QY 163 RSHLIADSLGDLARVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLVLYEVA 222  
Db 121 RSHLIADSLGDLARVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLVLYEVA 180  
QY 223 PYNADLIPRAVVMQSSDNTINEKLVYNTANGYTYNHGTPQK 271  
Db 181 PYNADLIPRAVVMQSSDNTINEKLVYNTANGYTYNHGTPQK 229

RESULT 7  
US-08-393-889-1  
; Sequence 1, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
US-08-393-889-1

Query Match 14.3%; Score 203; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14; Indels 0; Gaps 0;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNLGSRVFSKCRVLFVSMVALVSATMAVTTVLENTALAR 43  
Db 1 MNLGSRVFSKCRVLFVSMVALVSATMAVTTVLENTALAR 43

RESULT 8  
PCT-US94-09450-1  
; Sequence 1, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beckman Instruments, Inc.  
; STREET: 2500 Harbor Boulevard  
; CITY: Fullerton  
; STATE: California



COUNTRY: USA  
ZIP: 92634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
PCT-US94-09450-1

Query Match 14.3%; Score 203; DB 5; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MNLGSRVFSKKRLVKFMSVALVSATMAVTTVLTALAR 43  
Db 1 MNLGSRVFSKKRLVKFMSVALVSATMAVTTVLTALAR 43

RESULT 9  
US-08-393-889-6  
Sequence 6, Application US/08393889  
Patent No. 6420152  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-08-393-889-6

Query Match 14.2%; Score 202; DB 4; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 43 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLG 80  
Db 1 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLG 38

RESULT 10  
PCT-US94-09450-6  
Sequence 6, Application PC/TUS9409450  
GENERAL INFORMATION:  
APPLICANT: Beckman Instruments, Inc.  
APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes

PCT-US94-09450-6

Query Match 14.2%; Score 202; DB 5; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
|||||  
Db 1 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 38  
|||||

## RESULT 11

US-08-393-889-16  
; Sequence 16, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belai, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (819) 796-4000  
TELEFAX: (819) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes

US-08-393-889-16.

Query Match 12.0%; Score 170; DB 4; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.8e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 QTVQSNVDVLDGASKYLNEALAWTFNDSPNY 75  
|||||  
Db 1 QTVQSNVDVLDGASKYLNEALAWTFNDSPNY 32  
|||||

## RESULT 12

PCT-US94-09450-16  
; Sequence 16, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.

APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes

PCT-US94-09450-16

Query Match 12.0%; Score 170; DB 5; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.8e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 QTVQSNVDVLDGASKYLNEALAWTFNDSPNY 75  
|||||  
Db 1 QTVQSNVDVLDGASKYLNEALAWTFNDSPNY 32  
|||||

## RESULT 13

US-09-379-523-3  
; Sequence 3, Application US/09379523  
; Patent No. 6280993  
; GENERAL INFORMATION:  
; APPLICANT: YAMATO, Ichiro  
; APPLICANT: HOSAKA, Toshiaki  
; TITLE OF INVENTION: GENE ENCODING CLASS I COLLAGENASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burgess, Ryan and Wayne  
; STREET: 370 Lexington Avenue, Suite 2105  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 10017

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch disk  
COMPUTER: PC'S LIMITED SYSTEM 200  
OPERATING SYSTEM: WINDOWS 98  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:

```
/ APPLICATION NUMBER: US/09/379,523
/ FILING DATE: August 24, 1999
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Wayne, Milton J.
/ REGISTRATION NUMBER: 17,906
/ REFERENCE/DOCKET NUMBER: U-Wp-5462
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 212-683-8150
/ TELEFAX: 212-532-4285
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1118
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/
US-09-379-523-3

Query Match      8.1%; Score 114.5; DB 3; Length 1118;
Best Local Similarity 22.8%; Pred. No. 0.0055;
Matches 69; Conservative 45; Mismatches 100; Indels 89; Gaps 16;

QY 10 FSKKRLVFMVALVSATMAVTTVTLENTALARQTQVSNDDVVL-----NDGASK 59
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 13 YSKESKIQTVRRTVSLLAVLTMNTSSVLAKPIENTNDTSIKNVKLRNAPNEENSK 72
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 60 YLNEALAWTFNDSPNYKLTGTSQITPAPFKAGDILYKLDLGRTRTRAGFLTYANVE 119
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 73 KVBDSEK-----NDKVEHVK-----NIEEAKVEQAVEPKSKSTLRGASI- 111
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 120 GSYGVRSQFGKQNPAGWTGNPNHVYKIEWLNGLSV-----GDFWNR-SHLTADSL 171
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 112 -----ANTSEKYDFEYLNGLSYTELTLNLRIKNQINGLNFYST 152
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 172 G-----GDALRVNAVGTTRTONVGGK-----DQGGMRVTE-ORAEQWLEARNRDGVLVYE 220
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 153 GSQKFFGDKNRVQAINALQES--GRITYTANDMKGIETFEVLVLAGFYLYNDGLSYLN 210
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 221 VAPIYNADELIPRAVVSVQSSN-----TINEKV-----LVYN-TANGYTIYHNGTTP 268
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 211 DRNF--QDKIP--AMIAIQKNPFKLTAVQDEVITSLKLGIGNASANAENVN--NCVP 264
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 269 TQK 271
Db 265 VLK 267

RESULT 14
US-09-107-532A-5038
/ Sequence 5038, Application US/09107532A
/ Patent No. 6583275
/ GENERAL INFORMATION:
/ APPLICANT: Lynn A Doucette-Stamm and David Bush
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
/ ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
/
/ NUMBER OF SEQUENCES: 7310
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: GENOME THERAPEUTICS CORPORATION
/ STREET: 100 Beaver Street
/ CITY: Waltham
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02354
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: CD/ROM ISO9660
/ COMPUTER: PC
/ OPERATING SYSTEM: <Unknown>
/ SOFTWARE: ASCII
```

```
/
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/107,532A
/ FILING DATE: 30-Jun-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 60/085,598
/ FILING DATE: 14 May 1998
/ APPLICATION NUMBER: 60/051571
/ FILING DATE: July 2, 1997
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Ariniello, Pamela Deneke
/ REGISTRATION NUMBER: 40,489
/ REFERENCE/DOCKET NUMBER: GTC-012
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (781)893-5007
/ TELEFAX: (781)893-8277
/ INFORMATION FOR SEQ ID NO: 5038:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 97 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: YES
/ ORIGINAL SOURCE:
/ ORGANISM: Enterococcus faecium
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (B) LOCATION 1...97
/ SEQUENCE DESCRIPTION: SEQ ID NO: 5038:
/
US-09-107-532A-5038

Query Match      8.0%; Score 113; DB 4; Length 97;
Best Local Similarity 32.3%; Pred. No. 0.00018;
Matches 30; Conservative 14; Mismatches 41; Indels 8; Gaps 3;

QY 179 NAVGTTRTONVGGKQGGMRVTEQRAQWLEARNRDGVLVYEVAPIYNADELIPRAVVVS 238
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 9 NLMGTGRLNAPYM-----LAHENDIAVYIKETNH--HVRYRVTPEFEGNELVARGVQLE 61
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
QY 239 MQS-SDNTINEKVLVYNTANGYTIYHNGTPTQ 270
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 62 AESIEDKKIEFNVFIYVQDGYTINVTGQATK 94
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:

RESULT 15
US-08-393-889-4
/ Sequence 4, Application US/08393889
/ Patent No. 6420152
/ GENERAL INFORMATION:
/ APPLICANT: Adams, Craig W.
/ APPLICANT: Pang, Patty P.-Y.
/ APPLICANT: Belei, Marina
/ TITLE OF INVENTION: Recombinant DNase B Derived from
/ Streptococcus pyogenes
/ NUMBER OF SEQUENCES: 16
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Sheldon & Mak
/ STREET: 225 South Lake Avenue, Ninth Floor
/ CITY: Pasadena
/ STATE: California
/ COUNTRY: USA
/ ZIP: 91001
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/393,889
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/082,845
/ FILING DATE:
```

Search completed: January 5, 2004, 18:44:01  
Job time : 32.9429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:41:44 ; Search time 502.074 Seconds  
(without alignments)  
108.298 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MNLGSRVFSKCRLVKFS.....VNTANGYTYNHNGTPTQK 271

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 73397

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	271	9	US-09-119-900-8
2	1397.5	98.6	272	9	US-09-119-900-15
3	1220	86.0	229	9	US-09-119-900-9
4	203	14.3	43	9	US-09-119-900-1
5	202	14.2	38	9	US-09-119-900-6
6	170	12.0	32	9	US-09-119-900-16
7	131	9.2	261	12	US-09-769-736-129
8	126.5	8.9	274	12	US-09-769-744A-168
9	109.5	7.7	2386	15	US-10-156-761-7751
10	106	7.5	23	9	US-09-119-900-4
11	106	7.5	1233	10	US-09-738-626-4312
12	98	6.9	522	12	US-10-172-502-17
13	94	6.6	621	15	US-10-124-880-2
14	92.5	6.5	388	15	US-10-156-761-14080
15	92	6.5	234	12	US-10-441-626-3

16	92	6.5	234	14	US-10-075-872-3
17	92	6.5	234	15	US-10-261-997-3
18	92	6.5	867	10	US-09-839-894-6
19	90	6.3	398	12	US-10-369-493-354
20	90	6.3	2364	15	US-10-156-761-7834
21	89.5	6.3	275	12	US-10-323-069A-105
22	89	6.3	1232	12	US-10-369-493-21495
23	88.5	6.2	312	12	US-10-369-493-1039
24	88.5	6.2	596	12	US-10-238-075-416
25	87.5	6.2	608	15	US-10-156-761-8616
26	87	6.1	226	12	US-10-029-386-34033
27	87	6.1	362	12	US-10-369-493-786
28	87	6.1	422	12	US-10-369-493-1312
29	87	6.1	422	12	US-10-369-493-20377
30	87	6.1	428	12	US-10-284-400-8
31	87	6.1	429	12	US-10-156-761-9396
32	87	6.1	624	15	US-10-284-400-16
33	86.5	6.1	339	12	US-10-369-493-18526
34	86	6.1	1136	9	US-09-815-242-12447
35	86	6.1	1136	9	US-09-815-242-12814
36	86	6.1	1179	9	US-09-815-242-5522
37	86	6.1	1391	15	US-10-080-505-11
38	86	6.1	1391	15	US-10-080-505-15
39	85.5	6.0	1073	12	US-10-193-764-45
40	85.5	6.0	1079	12	US-10-193-764-43
41	85.5	6.0	1475	9	US-09-740-274-2
42	85.5	6.0	5636	12	US-10-032-189-128
43	85.5	6.0	5636	12	US-10-120-801-72
44	85.5	6.0	5636	12	US-10-023-634-93
45	85	6.0	420	12	US-10-369-493-21586

ALIGNMENTS

RESULT 1  
US-09-119-900-8  
; Sequence 8, Application US/09119900  
; Patent No. US20020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-119-900-8

Query Match      100.0%; Score 1418; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 9.9e-137;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVLTALARQTVSNVNDVNDGASKY 60
Db 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVLTALARQTVSNVNDVNDGASKY 60
QY 61 LNEALWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
Db 61 LNEALWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
QY 121 SYGVROSGFGKQNPAGTGNPNHVYKIEWLNGLSYVGFNWRSHLIADSLGDDALRVNA 180
Db 121 SYGVROSGFGKQNPAGTGNPNHVYKIEWLNGLSYVGFNWRSHLIADSLGDDALRVNA 180
QY 181 VTGRTQNVGGRDQKGMRYTEORAQEWLEARNRDGYLYVEVAPIYNADELIPRAVVVSMQ 240
Db 181 VTGRTQNVGGRDQKGMRYTEORAQEWLEARNRDGYLYVEVAPIYNADELIPRAVVVSMQ 240
QY 241 SSDNTINEKVLVNTANGYTIYHNGTPTQK 271
Db 241 SSDNTINEKVLVNTANGYTIYHNGTPTQK 271

RESULT 2
US-09-119-900-15
; Sequence 15, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/09/119,900
; APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/082,845
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-119-900-15

Query Match      98.6%; Score 1397.5; DB 9; Length 272;
Best Local Similarity 99.6%; Pred. No. 1.3e-134;
Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 NLLGSRVFSKCRVLKFSMVALVSATMAVTTVLTALARQTVSNVNDVNDGASKYL 61
Db 4 NLLGSRVFSKCRVLKFSMVALVSATMAVTTVLTALARQTVSNVNDVNDGASKYL 62
QY 62 NEALWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS 121
Db 63 NEALWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS 122
QY 122 YGVROSGFGKQNPAGTGNPNHVYKIEWLNGLSYVGFNWRSHLIADSLGDDALRVNAV 181
Db 123 YGVROSGFGKQNPAGTGNPNHVYKIEWLNGLSYVGFNWRSHLIADSLGDDALRVNAV 182
QY 182 TGRTQNVGGRDQKGMRYTEORAQEWLEARNRDGYLYVEVAPIYNADELIPRAVVVSMOS 241
Db 183 TGRTQNVGGRDQKGMRYTEORAQEWLEARNRDGYLYVEVAPIYNADELIPRAVVVSMOS 242
QY 242 SDNTINEKVLVNTANGYTIYHNGTPTQK 271
Db 243 SDNTINEKVLVNTANGYTIYHNGTPTQK 272

RESULT 3
US-09-119-900-9
; Sequence 9, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER: US/09/119,900
; APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/082,845
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
```

ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-9

Query Match 86.0%; Score 1220; DB 9; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.5e-116; Indels 0; Gaps 0;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 43 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALPKAGDILYSKLE 102  
Db 1 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALPKAGDILYSKLE 60  
Qy 103 LGRTRTARGTLTYANVSGSYGVROSGFNQNPAGTGNPNHVKYKIEWNLGSLSYVGDFWN 162  
Db 61 LGRTRTARGTLTYANVSGSYGVROSGFNQNPAGTGNPNHVKYKIEWNLGSLSYVGDFWN 120  
Qy 163 RSHLIADSLGDLARVNAVVTGTRTQNVGGDRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 222  
Db 121 RSHLIADSLGDLARVNAVVTGTRTQNVGGDRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 180  
Qy 223 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTTANGTYTINYHNGTPTQK 271  
Db 181 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTTANGTYTINYHNGTPTQK 229

## RESULT 4

US-09-119-900-1  
Sequence 1, Application US/09119900  
Patent No. US20020081622A1  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 795-6321  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-1

Query Match 14.3%; Score 203; DB 9; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.6e-13;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVRFVSKCRLVKFSMVALVSATMAVTTVTTLENTALAR 43  
Db 1 MNLGSRVRFVSKCRLVKFSMVALVSATMAVTTVTTLENTALAR 43

## RESULT 5

US-09-119-900-6  
Sequence 6, Application US/09119900  
Patent No. US20020081622A1  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 795-6321  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-6

Query Match 14.2%; Score 202; DB 9; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.7e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLG 80  
Db 1 RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLG 38

## RESULT 6

US-09-119-900-16  
Sequence 16, Application US/09119900  
Patent No. US20020081622A1  
GENERAL INFORMATION:

APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belel, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-16

Query Match 12.0%; Score 170; DB 9; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.5e-10;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 QTVSNDVVLNDGASKYLNEALAWTFNDSPNY 75  
DB 1 QTVSNDVVLNDGASKYLNEALAWTFNDSPNY 32

## RESULT 7

US-09-769-736-129  
Sequence 129, Application US/09769736  
Publication No. US20030138775A1  
GENERAL INFORMATION:  
APPLICANT: Microbial Technics Limited  
APPLICANT: Le Page, Richard WF  
APPLICANT: Wells, Jeremy M  
APPLICANT: Hanniffy, Sean B  
TITLE OF INVENTION: Proteins  
CURRENT APPLICATION NUMBER: US/09/769,736  
CURRENT FILING DATE: 2003-02-14  
PRIOR APPLICATION NUMBER: GB 9816335.5  
PRIOR FILING DATE: 1998-07-27  
PRIOR APPLICATION NUMBER: US 60/125163  
PRIOR FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 212  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 129  
LENGTH: 261

TYPE: PRT  
ORGANISM: Streptococcus agalactiae  
US-09-769-736-129  
Query Match 9.2%; Score 131; DB 12; Length 261;  
Best Local Similarity 26.0%; Pred. No. 5.6e-05;  
Matches 39; Conservative 25; Mismatches 60; Indels 26; Gaps 5;  
QY 131 NONPAGW--TGNPNHVYKIEWLNGLSYVGDFWNRSHLIADSLGCD-----ALRVNA 180  
DB 125 NWKPLGWHQVATNDHYCHAVD-----KGLHIAIYALAGNFKGWDASVSNPQNV 171  
QY 181 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVAPIYNAD-ELIPRAVVVSM 239  
DB 172 VTQTAHNSQSNQKINRGQNYYESILVRKAVDQNK--RVRYRVTPLYRNDTDLVFFAMHLEA 229  
QY 240 QSSDNTINEKVLVYNTANGYTYNHNGTPT 269  
DB 230 KSQDGLFEFNVAFPNTQASVTMDYATGEIT 259

## RESULT 8

US-09-769-744A-168  
Sequence 168, Application US/09769744A  
Publication No. US20030134407A1  
GENERAL INFORMATION:  
APPLICANT: Le Page, Richard WF  
APPLICANT: Wells, Jeremy M  
APPLICANT: Hanniffy, Sean B  
APPLICANT: Hansbro, Philip M  
TITLE OF INVENTION: Proteins  
FILE REFERENCE: PWC/P21122WO  
CURRENT APPLICATION NUMBER: US/09/769,744A  
CURRENT FILING DATE: 2001-01-26  
PRIOR APPLICATION NUMBER: PCT/GB99/02452  
PRIOR FILING DATE: 1999-07-27  
PRIOR APPLICATION NUMBER: GB 9816336.3  
PRIOR FILING DATE: 1998-07-27  
PRIOR APPLICATION NUMBER: US 60/125329  
PRIOR FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 196  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 168  
LENGTH: 274  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-769-744A-168

Query Match 8.9%; Score 126.5; DB 12; Length 274;  
Best Local Similarity 21.9%; Pred. No. 0.00017;  
Matches 59; Conservative 37; Mismatches 108; Indels 65; Gaps 11;

QY 32 TTVTLNTALARTQVSNVDVVLNDGASKYLNEALAWTFND-----GGDALR---VNAV 181  
DB 37 TNLQKQKQSEAPSAEASVLTDAVKSQIKGSLWNGSGAFVNGKNTLDAKVSSKPY 96  
QY 72 SPNYKTLTGTSQITPALPPKAGDILYKLDLDE--LGRTRTARGTLTYANVEGSGVROSGFG 129  
DB 97 ADNKTIVGKETV-----PTVANALLSKATQYKNRKETGNGSTSW-----137  
QY 130 KNQNPAGWTGNPNHVYKIEWLNGLSYVGDFWNRSHLIADSL-----GGDALR---VNAV 181  
DB 138 ---TPPGW-----HQVKNLKG-SYTHAV-DRGHLGLGALIGLDFDASTSNPKNIA 184  
QY 182 TGTRTONVGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVAPIYNAD-ELIPRAVVVSMQ 240  
DB 185 VQTAWANQAQAEYSTGQNYYESKVRKALDQNK--RVRYRVTPLYYASNEDLVPSASQIEAK 242  
QY 241 SSDNTINEKVLVYNTANGYTYNHNGTPT 269  
DB 243 SSDGELFEFNVLPNVQKGLQDLYRTGEVT 271



```

RESULT 10
US-09-119-900-4
; Sequence 4, Application US/09119900
; Patent No. US20020081632A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Beale, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
;

```

Query Match 7.5%; Score 106; DB 10; Length 1233;  
Best Local Similarity 22.8%; Pred. No. 0.2;  
Matches 55; Conservative 33; Mismatches 93; Indels 60



```
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14080

Query Match          6.5%; Score 92.5; DB 15; Length 388;
Best Local Similarity 23.7%; Pred. No. 0.88;
Matches 78; Conservative 32; Mismatches 114; Indels 105; Gaps 18;

QY 4 LGSRRVFSKKRLVKSFMVALVSATMAVTTVLTENTALARQTQVSNDDVVLNDGASKYLNE 63
DB 29 LKHRRISRRRAAVAGAGIAALVAAGVTFQTA---NASEAPKTEAPHTLSLS-AAGK---- 80
QY 64 ALAWTFNDSPNYKTLGTQITPALPPKAGDILYSKLDLGRTRTARGTLTYANVEGSYG 123
DB 81 -LASTLG-----KOLGTDAAAGTYIDAKAKHLVVNLDE-----TAAKTVEAAGAKARV- 127
QY 124 VRQSPGKNQNP-----AGWTGNPNHVKY-----KIEWLNGLSYVGDFWN 162
DB 128 VRNSLAELTSARTTLKQDATIPGTSWATDPETNKVVVTTADRTVSKAEWATLTKVVDGLGQ 187
QY 163 RSHL-----IADSLGGDALRVNAVGTGRTQNVGGRDQ-----KGGMRY----- 200
DB 188 RAELORTKGEYKFFIA---GGD-----AITGG-----GGRCSLGFNVVKGQPYFITAGH 234
QY 201 -TEQ-----RAQWLEARNRDGYLYYEVAPIYNAD-----ELIPRAVVVSMQ 240
DB 235 CTESISTWSDSGSGIQGTNEQSFPCNDFGLVKY-----TSNADHPSEVDLYNGSTQPI TK 290
QY 241 SSDNTINEKVLVYNTANGTYTINYHNGTPT 269
DB 291 AGDATVQKV-----TRSGSTTQVHSGTPT 315

RESULT 15
US-10-441-626-3
; Sequence 3, Application US/10441626
; Publication NO. US20030186418A1
; GENERAL INFORMATION:
; APPLICANT: Gualfetti, Peter
; APPLICANT: Mitchinson, Colin
; APPLICANT: Phillips, Jay Ian
; TITLE OF INVENTION: No. US20030186418A1el Variant EGIII-Like Cellulase
; TITLE OF INVENTION: Compositions
; FILE REFERENCE: GC631
; CURRENT APPLICATION NUMBER: US/10/441,626
; CURRENT FILING DATE: 2003-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Trichoderma reesei
US-10-441-626-3

Query Match          6.5%; Score 92; DB 12; Length 234;
Best Local Similarity 23.7%; Pred. No. 0.47;
Matches 70; Conservative 24; Mismatches 24; Indels 124; Gaps 17;

QY 17 VKFSMV--ALVSATMAVTTVLTENTALARQTQVSNDD-----VVLNDGASK 59
DB 1 MKFLQVLPALIPALAQAQTSQDQWATFTGNGYTVSNLWLGASAGSGFCVTAVSLSGASW 60
QY 60 YLNEALAWTFNDSPNYKTLGTQITPALPPKAGDILYSKLDLGRTRTAR 110
DB 61 HAD-----QWMSGQNNVKSQNSQIAIPQKRTVNSISSMPTTASWSYS----- 104
QY 111 GTLYANVEGSYGVRSQSGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDF-----WNRSHL 166
DB 105 GSNIRANV--AYDLFTA-----ANPNHV-----TYSGDYELMIW----- 136
QY 167 IADSLG--GDALRVNAVGTGRTQNVGGRDQKGGMRYTEQRAQEWLEARNRDGYLYYEVAPI 224
```

```
Db 137 ----LGKYGDIGPISQSG--TVNVCG-----QSWT-----LYYG----- 165
QY 225 YNADELIPRAVVVSMQSSDNTINEKVLV-----YNTANGYTTINYHNGT 267
DB 166 YNG-----AMQVVSFVAQTNTTNYSGDVKNFFNYLRDNKGYNAAAGQYVLSYQFGT 215

Search completed: January 5, 2004, 19:13:46
Job time : 505.074 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 17:31:14 ; Search time 31.442 Seconds  
(without alignments)  
828.882 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MNLGSRVFSKRLVKFS.....VYNTANGYTYNHNGTPTQK 271

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1414	99.7	271	2 J36907	mitogenic factor,
2	385	27.2	327	2 J70594	deoxyribonuclease
3	153.5	10.8	263	2 D86644	hypothetical prote
4	126.5	8.9	274	2 F95229	DNA-entry nuclease
5	126.5	8.9	274	2 S10641	endA protein - Str
6	126.5	8.9	274	2 A99094	deoxyribonuclease
7	104.5	7.4	455	2 G71896	probable outer mem
8	104	7.3	392	2 A56123	streptogrisin D (E
9	100	7.1	928	2 B84483	hypothetical prote
10	98.5	6.9	347	2 F97223	uncharacterized con
11	96.5	6.8	386	2 AB1491	conserved hypothet
12	95.5	6.7	319	2 S69209	alpha-toxin precu
13	95.5	6.7	839	2 E83521	probable fibrillar
14	94	6.6	595	2 B48658	flagellin - Escher
15	93.5	6.6	676	2 H96970	endo-arabinase rel
16	93	6.6	1203	2 S27545	pullulanase - Ther
17	93	6.6	1534	2 T30295	P-glycoprotein - T
18	92.5	6.5	482	2 E70460	flagellar hook pro
19	92	6.5	363	2 E84280	hypothetical prote
20	91.5	6.5	319	2 F89887	Alpha-Hemolysin pr
21	91.5	6.5	611	2 S76211	hypothetical prote
22	91.5	6.5	869	2 C56617	cfac protein precu
23	91	6.4	638	2 D69957	conserved hypothet
24	90.5	6.4	292	2 S58857	botulinum neurotox
25	90.5	6.4	4199	2 S76412	hypothetical prote
26	90	6.3	386	2 AG1120	conserved hypothet
27	90	6.3	474	2 T31104	hemolysin accessor
28	89.5	6.3	743	2 AH2060	hypothetical prote
29	89	6.3	396	2 JC5558	alpha-galactosidas

30	89	6.3	488	2 A11930	proteinase [import
31	89	6.3	1232	2 D64413	cobalamin biosynth
32	88.5	6.2	255	2 AE3449	periplasmic mannit
33	88.5	6.2	312	2 F64435	mevalonate kinase
34	88	6.2	362	2 D90755	outer membrane pro
35	88	6.2	362	2 B85619	outer membrane pro
36	88	6.2	646	2 H95155	prolyl oligopeptid
37	88	6.2	646	2 C98022	oligopeptidase B (
38	88	6.2	774	2 AG1565	autolysin (amidase
39	88	6.2	802	2 I39665	penicillin amidase
40	88	6.2	1162	2 T30433	scaffolding protei
41	87.5	6.2	941	2 S29043	cellulase (EC 3.2.
42	87.5	6.2	2154	2 F83068	hypothetical prote
43	87	6.1	362	1 MMECF	outer membrane por
44	87	6.1	420	1 JN0854	glutamate dehydrog
45	87	6.1	422	2 A71038	probable glutamate

ALIGNMENTS

RESULT 1

S36907

mitogenic factor, 25K, precursor - Streptococcus pyogenes

C;Species: Streptococcus pyogenes

C;Date: 10-Dec-1993 #sequence,revision 23-Feb-1996 #text\_change 15-Oct-1999

C;Accession: S36907; S36908; S29188

R;Iwasaki, M.; Igarashi, H.; Hinuma, Y.; Yutsudo, T.

FEBS Lett. 331, 187-192, 1993

A;Title: Cloning, characterization and overexpression of a Streptococcus pyogenes gene e

A;Reference number: S36907; MUID:94009636; PMID:8405402

A;Accession: S36907

A;Molecule type: DNA

A;Residues: 1-271 <IWA>

A;Cross-references: EMBL:DJ3428; NID:g432369; PIDN:BAA02693.1; PID:d1003198; PID:g432370

A;Accession: S36908

A;Molecule type: protein

A;Residues: 44-64 <IWW>

R;Yutsudo, T.; Murai, H.; Gonzalez, J.; Takao, T.; Shimonishi, Y.; Takeda, Y.; Igarashi,

FEBS Lett. 308, 30-34, 1992

A;Title: A new type of mitogenic factor produced by Streptococcus pyogenes.

A;Reference number: S29188; MUID:92354761; PMID:1644200

A;Accession: S29188

A;Molecule type: protein

A;Residues: 44-64 <YUT>

F;1-43/Domain: signal sequence #status predicted <SIG>

F;44-271/Product: mitogenic factor, 25K #status experimental <MAT>

Query Match 99.7%; Score 1414; DB 2; Length 271;  
Best Local Similarity 99.6%; Pred. No. 9.9e-111;  
Matches 270; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MNLGSRVFSKRLVKFSMVALVSNATWATVTTLENTALARQTQVSNVDVNDGASKY	60
Db	1	MNLGSRVFSKRLVKFSMVALVSNATWATVTTLENTALARQTQVSNVDVNDGASKY	60
Qy	61	LNELAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG	120
Db	61	LNELAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG	120
Qy	121	SYGVRSFGKNQNPAGWTGNPNHVKYKIEWNLGSLVYGVDFWNRSHLIADSLGGDALRVNA	180
Db	121	SYGVRSFGKNQNPAGWTGNPNHVKYKIEWNLGSLVYGVDFWNRSHLIADSLGGDALRVNA	180
Qy	181	VTGTRTNVGGRRQKGMRYTEQRAQEWLEARNRDGILYIEVAPIYNADELIPRAVYVSMQ	240
Db	181	VTGTRTNVGGRRQKGMRYTEQRAQEWLEARNRDGILYIEVAPIYNADELIPRAVYVSMQ	240
Qy	241	SSDNTINEKVLVNTANGYTYNHNGTPTQK	271
Db	241	SSDNTINEKVLVNTANGYTYNHNGTPTQK	271

## RESULT 2

JT0584  
deoxyribonuclease I (EC 3.1.21.1) sdc precursor - Streptococcus "equisimilis"  
N:Alternate names: streptodornase  
C:Species: Streptococcus "equisimilis"  
C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 15-Oct-1999  
C:Accession: JT0584; S24204  
R:Wolynowska, R.; Ceglowski, P.; Kok, J.; Venema, G.  
Gene 106, 115-119, 1991  
A:Title: Isolation, sequence and expression in *Escherichia coli*, *Bacillus subtilis* and *Lactococcus lactis*  
A:Reference number: JT0584; MUID:92039051; PMID:1937032  
A:Accession: JT0584  
A:Molecule type: DNA  
A:Residues: 1-327 <WOL>  
A:Cross-references: EMBL:X17241; NID:948693; PIDN:CAA35106.1; PID:948694  
A:Experimental source: strain H46A  
C:Genetics:  
A:Gene: sdc  
C:Keywords: hydrolase  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:36-308/Product: deoxyribonuclease sdc #status predicted <MAT>

Query Match 27.28; Score 385; DB 2; Length 327;  
Best Local Similarity 33.08; Pred. No. 1.4e-24;  
Matches 100; Conservative 39; Mismatches 96; Indels 68; Gaps 8;  
QY 11 SKCR--LVKFSVALVS-ATMAVTTVLTENTALARQTOVNDVLDGASKYLNEALAW 67  
DB 2 SKLRNLFVRIIAAFASPAVMAIPVH-HNTVLAKTVSVNQ----- 42  
QY 68 TENDSPNYKTLGTSQITPALPK-----AGDILYSKLDLGRTR 107  
DB 43 TYGEYKDYTVIGESNIDQSAPFKYIKTTERVYKQGTSEKRVTVSDVVYNELDYKKEST 102  
QY 108 TARGTLTYANVEGSGVROSGNQNPAWG-----TGN-----PNH 143  
DB 103 GAVGVVTKMDMSKGYREKWEINPEPSGFYFYNRANDEEISEKEYSRRTSKYKVTNN 162  
QY 144 VKYKIEWNLGSLVGVDFWNRSHLIADSLGGDALRVNAVGTTRQNVGGDKGGMRYTEQ 203  
DB 163 VPVLVTLTKGKYNHLFVASHLFADSLGGKIRKNAITGTQMNQVGR--KGGQVYIEK 220  
QY 204 RAQEWLEARNRGVLYVEVAPIYNADELIPRAVVSQSDNTINEKLVVYNTANGYTYNY 263  
DB 221 KVLSHITKPNVDVYFYSALPEYQGAELLARSLVLSALSDDGVINETVRVNTADGFNIN 280  
QY 264 HNG 266  
DB 281 ERG 283

## RESULT 3

D86644  
hypothetical protein ybfB (imported) - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: D86644  
R:Botolin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: D86644  
A:Molecule type: DNA  
A:Status: preliminary  
A:Residues: 1-263 <STO>  
A:Cross-references: GB:AB005176; PID:g12723006; PIDN:AAK04254.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: ybfB

Query Match 10.88; Score 153.5; DB 2; Length 263;  
Best Local Similarity 24.48; Pred. No. 2.3e-05;  
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;

QY 68 TENDSPNYKTLGTS-QITPALPKAGDILYSKLDLGR-----RTARGTL 113  
DB 39 TDNSQVSTKSLASSVKQAPLTFKNQRMVNMANTDALGRAVDHSIQLKDSQEPKVKREPL 98  
QY 114 TVANVEGSGVROSGFKQKNPAGWTGNPNVYKIEWNLGSLVYGVDFW--NRSHLIA--- 168  
DB 99 TV-----NPVGW-----H-NYNFYKSDSGSKGKWLWARGHLVGYQF 135  
QY 169 DSLGGDALRV-----NAVGTTRQNVGGDKGGMRYTEQRAQEWLEARNRGDGLYY 219  
DB 136 SCLNNEARNLVPEFAWFGNGNFTGTNDGNT-----ASMLYYENRLDSWLANHPNYLDY 189  
QY 220 EVAPIYNADELIPRAV--VVSQSDNTINEK-----VLVYNTANGYTY 261  
DB 190 QVTPLYEGNELLPRQIRLAYVGDKNQQTLSIKLGGREKSGNGGATVVVLDNVAPNAKI 249  
QY 262 NYHNGT 267  
DB 250 NYADGT 255

## RESULT 4

F95229  
DNA-entry nuclease (imported) - Streptococcus pneumoniae (strain TIGR4)

C:Species: Streptococcus pneumoniae

C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001

C:Accession: F95229

R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heide

son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel, I.

Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A:Title: Complete Genome Sequence of a virulent isolate of *Streptococcus pneumoniae*.

A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: F95229

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-274 <KUR>

A:Cross-references: GB:AE005672; PIDN:AAK76031.1; PID:g14973470; GSPDB:GN00164; TIGR:SP45

A:Experimental source: strain TIGR4

C:Genetics:

A:Gene: SP1964

Query Match 8.98; Score 126.5; DB 2; Length 274;  
Best Local Similarity 21.98; Pred. No. 0.0044;  
Matches 59; Conservative 37; Mismatches 108; Indels 65; Gaps 11;

QY 32 TTVLTENTALARQTOVNDVLDGASKYLNEALAWTEND----- 71  
DB 37 TNLQKQKQASEAPSALESVLTDAVKQIKGSLFWNGSGAFVNGNKTNLDAKVSSKPY 96

QY 72 SPNYKTLGTSQITPALPKAGDILYSKLDL--LGRTRTARGTLTYANVEGSGVROSGFG 129  
DB 97 ADNKTKTGVKETV-----PTVANALLSKATQYKRNKRTGNGSTSW----- 137

QY 130 KNQNPAGWTGNPNVYKIEWNLGSLVYGVDFWNRSHLIADSL-----GGDALR---VNAV 181  
DB 138 ---TPGW-----HQQNKLKG-SYTHAV-DRGHLLGVALLGLDGFDASTSNPKNIA 184

QY 182 TGTTRQNVGGDKGGMRYTEQRAQEWLEARNRGVLYVEVAPIYNADELIPRAVVSQ 240  
DB 185 VQTAWANQAQAEYSTGQNYYESKYRKALDQNK--RVRYRTLYYASNEDLVPSASQIEAK 242

QY 241 SSDNTINEKLVVYNTANGYTYNYHNGTPT 269  
DB 243 SSDGELFBNLVFPNVQKGLQDYRTGEVT 271

## RESULT 5

S10641

endA protein - Streptococcus pneumoniae

C:Species: Streptococcus pneumoniae

C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 15-Oct-1999

C;Accession: S10641  
R;Puyet, A.; Greenberg, B.; Lacks, S.A.  
J. Mol. Biol. 213, 727-738, 1990  
A;Title: Genetic and structural characterization of endA. A membrane-bound nuclease required for DNA replication in *Escherichia coli*. PMID:2359120  
A;Reference number: S10640; MUID:90294291; PMID:2359120  
A;Accession: S10641  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-274 <PUY>  
A;Cross-references: GB:X54225; NID:g47372; PIDN:CAA38134.1; PID:g47374

Query Match 8.9%; Score 126.5; DB 2; Length 274;  
Best Local Similarity 21.9%; Pred. No. 0.0044;  
Matches 59; Conservative 37; Mismatches 108; Indels 65; Gaps 11;  
  
Qy 32 TTVTLENTALARQTVSNVDVNDGASKYLEALAWTFND----- 71  
Db 37 TNLQKQASAPQALAESVLTDAVKSQIKSLEWNGSGAFVNGKTNLDKAVSSKPY 96  
Qy 72 SPNYKTLGTSQITPALPPKAGDILYKLDL--LGRTRTARGTLTYANVEGSGVROSPG 129  
Db 97 ADNKTKTGKETV-----PTVANALLSKATQYKRNKGTGSGTSM----- 137  
Qy 130 KNQNPAGWTGNPNHVKYKIEWLNGLSYVGDVFNRSHLIADSL-----GGDALR---VNAV 181  
Db 138 ---TPPGW-----HQQVKNLKG-SYTHAV-DRGHLLGYALIGLGDGFDASTSNPKNIA 184  
Qy 182 TGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYVEVAPIYNADE-LIPRAVVSMQ 240  
Db 185 VQTAWANQAQAEYSTGQNYYESKVRKALDQNK--RVRYRVTLYYASNEDLVPSASQIEAK 242  
Qy 241 SSDNTINEKVLVYNTANGTYTINYHNGTPT 269  
Db 243 SSDGELEFNVLPVNVQKGLDLYRTGEVT 271

## RESULT 6

A99094  
deoxyribonuclease I (EC 3.1.30.-) [imported] - Streptococcus pneumoniae (strain R6)  
C;Species: Streptococcus pneumoniae  
C;Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C;Accession: A99094  
R;Hoekings, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; H  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A;Reference number: A97872; MUID:21429245; PMID:11544234  
A;Accession: A99094  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-274 <KUR>  
A;Cross-references: GB:AB007317; PIDN:AAL00582.1; PID:g15459463; GSPDB:GN00174  
C;Genetics:  
A;Gene: endA

Query Match 8.9%; Score 126.5; DB 2; Length 274;  
Best Local Similarity 21.9%; Pred. No. 0.0044;  
Matches 59; Conservative 37; Mismatches 108; Indels 65; Gaps 11;  
  
Qy 32 TTVTLENTALARQTVSNVDVNDGASKYLEALAWTFND----- 71  
Db 37 TNLQKQASAPQALAESVLTDAVKSQIKSLEWNGSGAFVNGKTNLDKAVSSKPY 96  
Qy 72 SPNYKTLGTSQITPALPPKAGDILYKLDL--LGRTRTARGTLTYANVEGSGVROSPG 129  
Db 97 ADNKTKTGKETV-----PTVANALLSKATQYKRNKGTGSGTSM----- 137  
Qy 130 KNQNPAGWTGNPNHVKYKIEWLNGLSYVGDVFNRSHLIADSL-----GGDALR---VNAV 181  
Db 138 ---TPPGW-----HQQVKNLKG-SYTHAV-DRGHLLGYALIGLGDGFDASTSNPKNIA 184

Qy 182 TGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYVEVAPIYNADE-LIPRAVVSMQ 240  
Db 185 VQTAWANQAQAEYSTGQNYYESKVRKALDQNK--RVRYRVTLYYASNEDLVPSASQIEAK 242  
Qy 241 SSDNTINEKVLVYNTANGTYTINYHNGTPT 269  
Db 243 SSDGELEFNVLPVNVQKGLDLYRTGEVT 271

## RESULT 7

G71896  
Probable outer membrane protein - Helicobacter pylori (strain J99)  
C;Species: Helicobacter pylori  
A;Variety: strain J99  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 26-May-2000  
C;Accession: G71896  
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;  
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;  
Nature 397, 176-180, 1999  
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path  
A;Reference number: A71800; MUID:99120557; PMID:9923682  
A;Accession: G71896  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-455 <ARN>  
A;Cross-references: GB:AE001503; GB:AE001439; NID:g4155275; PIDN:AAD06302.1; PID:g415528  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: jhp0719  
C;Superfamily: Helicobacter pylori hypothetical protein HP0209

Query Match 7.4%; Score 104.5; DB 2; Length 455;  
Best Local Similarity 25.2%; Pred. No. 0.6;  
Matches 53; Conservative 19; Mismatches 85; Indels 53; Gaps 8;  
  
Qy 71 DSPNYKTLGTSQITP--ALFPKAGDILY-----SKLDLGRTRTARGTLTYANVEGSY 122  
Db 261 DSNPKFKGLGLRAQTITNVIFVYAKDLYDVYWRNSKIGEWGASLLIHQRFDYNEFNFGF 320  
Qy 123 GVRQSGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDVFNRSHLIADSLGDALRVNAV 182  
Db 321 GYQNFQGNANARIGWYGNPIPFNRY-----NNSVYGGVFSN-----AITADAVS 364  
Qy 183 GTRTQNVGGRDQKGM-----RYT-EQRAQEWLEARNRDLGYVEVAPIYNADELIPRAVV 236  
Db 365 G---YVFGGQVYRGFLWGILGRITYATRASERINLNLYGKWSFARV----- 409  
Qy 237 VMSQSDNTINEKVLVYNTANGTYTINYHNG 266  
Db 410 -----DVNLEYVYVVMHNGYRLDLYLTG 431

## RESULT 8

A56123  
streptogrisin D (EC 3.4.21.-) precursor - Streptomyces griseus (strain IMRU3499)  
C;Species: Streptomyces griseus  
C;Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 16-Aug-2002  
C;Accession: A56123  
R;Sidhu, S.S.; Kaimar, G.B.; Willis, L.G.; Borgford, T.J.  
J. Biol. Chem. 270, 7594-7600, 1995  
A;Title: Protease evolution in Streptomyces griseus. Discovery of a novel dimeric enzyme  
A;Reference number: A56123; MUID:95221424; PMID:7706307  
A;Accession: A56123  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-392 <SID>  
A;Cross-references: GB:IL29019; NID:g755072; PIDN:AAA74409.1; PID:g755073  
C;Genetics:  
A;Gene: sprD  
A;Start codon: GTG  
C;Superfamily: streptogrisin A  
C;Keywords: hydrolase; serine proteinase







Db 285 EVNFDVDANGKITIGGQAYLTSDGNLTNDAGGA-----TAATLDGLFKAGDQGS 336  
QY 160 -FWRSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEW----- 208  
Db 337 IGFNKTASV--TWGGTT--YNFKTA--DAGAATANAGVSFTDTASKETVLNKVATAKQ 389  
QY 209 ----LEARNR-----GYLYYEVAPIYNADLIPRAVVVSMOSSDNT-INERKVLVYNT 255  
Db 390 GTAVAANGDTSATITTKSGVQTYQA--VFAAGD-----GTASAKYADNTDVSNATATYTD 442  
QY 256 ANG-----YTINYHNGTPT 269  
Db 443 ADGEMTTIGSYTYTKYSIDANNKVT 467

RESULT 15  
H96970  
endo-arabinase related enzyme (family 43 glycosyl hydrolase domain and ricin B-like domain)  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: H96970  
R:Noelling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum ATCC824  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: H96970  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-676 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK7855.1; PID:gl5023444; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC0577

Query Match 6.6%; Score 93.5; DB 2; Length 676;  
Best Local Similarity 21.7%; Pred. No. 8.5;  
Matches 64; Conservative 47; Mismatches 93; Indels 91; Gaps 17;  
QY 19 FSMVALVSA-----TWAVTTVLTENTALARQTVQVSNVVDVNDGASKYLNEALAW 67  
Db 137 YSSIALATSNITGPIKYAGTIVYSGFT--NSSLASQTDYSKVGTGTNNVASRYLSKG-AW 193  
QY 68 TFDNSPNYKTLGTSQITPAL-FPKAGDILYS-----KLDELGRTRTARGTL-T 114  
Db 194 NSSYGN-----AIDPCVKYKSGNLWLSYGSWFGGIFWLKLDKSTGLRDYSYTYST 245  
QY 115 YANVEGSI-GVRSQFGQNQNPAGWTGNPNHVKY-----IEWNLGLSYVGDVFNKSH 165  
Db 246 KTNASQYLGKISGG-----YGTGEGSYIVYDKATDYVYLYESYCGLD----- 290  
QY 166 LIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWL-----EARNR----- 214  
Db 291 -ATDNFSGVHRL-----FRSKNITG-----PYTDAGNPAICTSANDNKNKGKILF 337  
QY 215 GYLYYEVAPIYNADLIPRAVVVSMOSSD--NTINEKVLVYNTANGVTINYHNGT 267  
Db 338 GNYFSSLSVNGSELSSKGYMSGCHNSAIDTSGQRVLIYHT-----RPNNGT 386

Search completed: January 5, 2004, 18:42:51  
Job time : 36.442 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 15:46:12 ; Search time 22.9576 Seconds  
(without alignments)  
555.120 Million cell updates/sec

Title: US-08-482-785-8

Perfect score: 1418

Sequence: 1 MNLLGSRVPSKCLRVKES.....VYNTANGYINYNHGTPTQK 271

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	385	27.2	327	1 DRN1_STREQ	P26295 streptococ
2	126.5	8.9	274	1 NUCE_STRPN	Q03158 streptococ
3	104	7.3	322	1 PRTE_STRGR	F52321 streptomyc
4	100.5	7.1	1279	1 APU_THESA	P36905 t amylopul
5	99	7.0	1861	1 APU_THETU	P38536 t amylopul
6	95.5	6.7	319	1 HLA_STAAU	P09616 staphyloc
7	94	6.6	341	1 OMPL_PHOPR	Q52581 photobacter
8	92.5	6.5	482	1 FLGE_AQUAE	Q67711 aquifex aeo
9	91.5	6.5	869	1 CFAC_ECOLI	P25733 escherichia
10	91	6.4	638	1 YQGS_BACSU	P54496 bacillus su
11	88.5	6.2	312	1 KIME_METJA	Q58487 methanococ
12	88.5	6.2	1476	1 GTFB_STRMU	P08987 streptococ
13	88.5	6.2	2208	1 POLN_MANCV	Q69014 manchester
14	88	6.2	802	1 PAC_ARTVI	P31956 arthrobacte
15	87.5	6.2	941	1 GUN_BACS6	P19424 bacillus sp
16	87.5	6.2	1047	1 POL_SIVAI	P27973 simian immu
17	87	6.1	362	1 OMPE_ECOLI	P29331 escherichia
18	87	6.1	420	1 DHE3_PVRFU	P80319 pyrococcus
19	87	6.1	420	1 DHE3_PVRHO	Q52310 pyrococcus
20	86.5	6.1	1116	1 RPOB_HETCA	P36440 heterosigma
21	86.5	6.1	1709	1 CHDI_HUMAN	O14646 homo sapien
22	86	6.1	288	1 BSN2_BACSU	Q32150 bacillus su
23	86	6.1	296	1 NIFD_NOSCO	P52337 nostoc comm
24	86	6.1	1182	1 RPOB_STAAU	P47768 staphylococ
25	85	6.0	213	1 GL25_ARATH	Q65252 arabidopsis
26	85	6.0	420	1 DHE3_PPRAB	Q47950 pyrococcus
27	85	6.0	524	1 CK13_YEAST	P39962 saccharomyc
28	85	6.0	1165	1 RPC2_SCHPO	Q10233 schizosacch
29	84.5	6.0	1723	1 PM20_CHLPN	Q92812 chlamydia p
30	84	5.9	800	1 GUN_BACS1	P06564 bacillus sp
31	84	5.9	1102	1 RPOE_STNY3	P77965 synecocyst
32	84	5.9	1196	1 AMYB_PABPO	P21543 paenibacill
33	83.5	5.9	227	1 XYN1_HUMIN	P55334 humicola in

34	83.5	5.9	504	1 HUTH_STAAW	Q8ny73 staphylococ
35	83.5	5.9	856	1 AAP1_YEAST	P37896 saccharomyc
36	83.5	5.9	922	1 PMP1_CHLPN	Q92995 chlamydia p
37	83.5	5.9	973	1 PM13_CHLPN	Q92896 chlamydia p
38	83.5	5.9	1157	1 XYN1_THESA	P36917 thermoanaer
39	83.5	5.9	3093	1 POLG_BSTV1	Q65730 b genome po
40	83	5.9	398	1 KBL_ECOLI	P07912 escherichia
41	83	5.9	398	1 KBL_SALTY	P37419 salmonella
42	83	5.9	609	1 HAPT_VIBCH	P24153 vibrio chol
43	83	5.9	666	1 MUR2_ENTHR	P39046 enterococcu
44	83	5.9	817	1 KPRO_MAIZE	P17801 zea mays (m
45	82.5	5.8	480	1 NIFD_CYAA5	O07642 cyanothece

#### ALIGNMENTS

##### RESULT 1

ID	DRN1_STREQ	STANDARD;	PRT;	327 AA.
AC	P26295;			
DT	01-MAY-1992 (Rel. 22, Created)			
DT	01-MAY-1992 (Rel. 22, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	Deoxyribonuclease precursor (EC 3.1.21.1) (Streptodornase) (DNase).			
GN	SDC.			
OS	Streptococcus equisimilis.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCEI_TaxID=119602;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=H46A;			
RC	MEDLINE=92039051; PubMed=1937032;			
RA	Wolinowska R., Ceglowski P., Kok J., Venema G.;			
RT	"Isolation, sequence and expression in Escherichia coli, Bacillus			
RT	subtilis and Lactococcus lactis of the DnaE			
RT	(streptodornase)-encoding gene from Streptococcus equisimilis H46A.";			
RL	Gene 106:115-119(1991).			
CC	-!- FUNCTION: MAY HAVE A ROLE IN S.EQUISIMILIS VIRULENCE.			
CC	-!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-			
CC	phosphodinucleotide and 5'-phosphogluconucleotide end-products.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X17241; CAA35106.1; -			
DR	PIR; J0584; J0584.			
KW	Hydrolase; Endonuclease; Nuclease; Signal.			
FT	SIGNAL 1 24 OR 35 (POTENTIAL).			
FT	CHAIN 25 327 DEOXYRIBONUCLEASE.			
SQ	SEQUENCE 327 AA; 36844 MW; 7529702FDF8E4AF7 CRC64;			

Query Match 27.2%; Score 385; DB 1; Length 327;

Best Local Similarity 33.0%; Pred. No. 7.7e-26;

Matches 100; Conservative 39; Mismatches 96; Indels 68; Gaps 8;

QY 11 SKKCR--LVKFSNVALVS-ATWAVTTVLTALAPQTVSNVDVLDGASKYLNEALAW 67

Db 2 SKKLRNLFRIIVAAPASFAVMAIPPYH-HNTVLAKTVSNQ----- 42

QY 68 TFNDSPNYKTLTQTSQITPALFPK-----AGDILYSKLDELGRTR 107

Db 43 TYGEYKDYTVIGESNIDQSAFPKIYKTVYKGGTSEKRTVTVSDVYNPLDGYKRST 102

QY 108 TARGTLTYANVGSGYGVRSFGKNQNPAGW-----TGN-----PNH 143

Db 103 GAYGVVTKDMIDMSKGYREKWTNPSPGFRFPYNRADNEEISEKEYDSRRKSYKVTNN 162

144 VKYKIEWLNGSLVYGDFFWNRSHLIADSLGGDALRVNAVTRTQNYGGRDQKGMRYTEQ 203  
 163 VPVVLTTLTKGKKYNSHLFVASHLFGSLGKSKIRKNAITQTQONVQTR--KGMQVTEK 220  
 204 RAQEWLEANDRGVLYYEVAPINADELIPRAVVVSMOSSDNTINEKLVVYNTANGTYTNY 263  
 221 KVLSHYTNPDVYVYSAIPEYQAEALLARSLVLSLSDGVINETVRVNTADGFNINY 280  
 264 HNG 266  
 281 EKG 283

RESULT 2  
 NUCE STRPN STANDARD; PRT; 274 AA.  
 AC NUCE STRPN STANDARD; PRT; 274 AA.  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE DNA-entry nuclease (Competence-specific nuclease) (EC 3.1.30.-).  
 GN ENDA OR SPI1964 OR SPI1779.  
 OS Streptococcus pneumoniae, and  
 OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1313, 171101;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=470;  
 RX MEDLINE=90294291; PubMed=2359120;  
 RA Puyet A., Greenberg B., Lacks S.A.;  
 RT "Genetic and structural characterization of endA. A membrane-bound  
 RT nuclease required for transformation of Streptococcus pneumoniae.";  
 RL J. Mol. Biol. 213:727-738(1990).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-334 / TIGR4;  
 RX MEDLINE=21357209; PubMed=11544234;  
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
 RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,  
 RA Durkin A.S., Gwinn M., Kolonay J.P., Nelson W.C., Peterson J.D.,  
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
 RA Holtzapple E., Khouri H., Wolf A.M., Uterback T.R., Hansen C.L.,  
 RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickson T., Hickey E.K.,  
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
 RT "Complete genome sequence of a virulent isolate of Streptococcus  
 RT pneumoniae";  
 RL Science 293:498-506(2001).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC BAA-255 / R6;  
 RX MEDLINE=21429245; PubMed=11544234;  
 RA Hoshins J., Alborn W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,  
 RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,  
 RA Gilnour R., Glass J.S., Khaja H., Kraft A.R., Lagace R.E.,  
 RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsumura P.,  
 RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,  
 RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,  
 RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,  
 RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,  
 RA Glass J.I.;  
 RT "Genome of the bacterium Streptococcus pneumoniae strain R6";  
 RL J. Bacteriol. 183:5709-5717(2001).  
 CC -!- FUNCTION: By degrading DNA that enters the cell, plays a role in  
 CC the competence of cells to be transformed.  
 CC -!- SUBCELLULAR LOCATION: Membrane-bound.  
 CC -!- SIMILARITY: Belongs to the DNA/RNA non-specific endonuclease  
 CC family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X54225; CAA38134.1; -  
 CC EMBL; AB007486; AAK76031.1; -  
 CC EMBL; AE008543; AAL00582.1; -  
 CC PIR; A99094; A99094.  
 CC PIR; F95229; F95229.  
 CC PIR; S10641; S10641.  
 CC TIGR; SPI964; -  
 CC InterPro; IPR001604; Endonuclease.  
 CC Pfam; PF01223; Endonuclease; 1.  
 CC PROSITE; PS01070; NUCLEASE\_NON\_SPEC; 1.  
 CC Competence; Hydrolase; Nuclease; Endonuclease; Signal-anchor;  
 CC Transmembrane; Complete proteome.  
 CC TRANSMEM 8 25 SIGNAL-ANCHOR (POTENTIAL).  
 CC ACT SITE 160 160 BY SIMILARITY.  
 CC SEQUENCE 274 AA; 29890 MW; 59B2243F0150CD98 CRC64;  
 CC  
 CC Query Match 8.9%; Score 126.5; DB 1; Length 274;  
 CC Best Local Similarity 21.9%; Pred. No. 0.0011;  
 CC Matches 59; Conservative 37; Mismatches 108; Indels 65; Gaps 11;  
 CC  
 CC 32 TTVTLENTALAROTQVSDNVDLNDGASKYLNEALMTFND----- 71  
 CC 37 TNLGSKQASEAPSAQALAESVLTDVAKSQIKGLEWNGCAFVNGKTNLDKAVSSKPY 96  
 CC 72 SPNVYKTLGTSQITPALFPKAGDILYSKLDE--LGRTRTARGTLTYANVEGSGYVRQSGF 129  
 CC 97 ADNKTXTVGKTV-----PTVANALLSKATQYKRNKKTGTGSGTSW----- 137  
 CC 130 KNQNPAGWTGNPHVKYKIEWLNGSLVYGDFFWNRSHLIADSL-----GGDALR---VNAV 181  
 CC 138 ---TPPCW-----HOVKNLKG-SYTHAV-DRGHLGLYALGGDLGDFDASTSNPKNIA 184  
 CC 182 TGTRTQNVGGRDQKGMRYTEQRAQEWLEANDRGVLYYEVAPINADE-LIPRAVVVSMQ 240  
 CC 185 VQTAWANQAQAEYSTQNYTESKVRKALDNQK--RVRYRVTLYASNEDLVPSASQIEAK 242  
 CC 241 SSDNTINEKLVVYNTANGTYTNYHNGTPT 269  
 CC 243 SSDGELEFNVLPVNVOKGLQLDYRTGEVT 271  
 CC  
 CC RESULT 3  
 CC ID PRD STRGR STANDARD; PRT; 392 AA.  
 CC AC P52321;  
 CC DT 01-OCT-1996 (Rel. 34, Created)  
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Serine protease D precursor (EC 3.4.21.-) (SGPD).  
 CC GN SPED.  
 CC OS Streptomyces griseus.  
 CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 CC OC Streptomycineae; Streptomycetaceae; Streptomycetes.  
 CC OX NCBI\_TaxID=1911;  
 CC [1]  
 CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 CC RC STRAIN=IMRU 3499;  
 CC RX MEDLINE=95221424; PubMed=7706307;  
 CC RA Sidhu S.S., Kalmr G.B., Willis L.G., Borgford T.J.;  
 CC RT "Protease evolution in Streptomyces griseus. Discovery of a novel  
 CC dimeric enzymes.";  
 CC RL J. Biol. Chem. 270:7594-7600(1995).  
 CC -!- FUNCTION: HAS A PRIMARY SPECIFICITY FOR LARGE ALIPHATIC OR  
 CC AROMATIC AMINO ACIDS.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY SZA.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L29019; AAA74409.1; -
DR PIR; A56123; A56123.
DR HSSP; P41140; 2SFA.
DR MEROPS; S01.266; -.
DR InterPro; IPR004236; AL protease.
DR InterPro; IPR001316; Endopptdase2A.
DR InterPro; IPR001254; Ser protease_Try.
DR Pfam; PF02983; AL protease; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00861; ALTYICTPASE.
DR SMART; SM00200; Tryp_Spc; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 64
FT PROPEP 65 204
FT CHAIN 205 392
FT ACT_SITE 237 237
FT ACT_SITE 266 266
FT ACT_SITE 348 348
FT ACT_SITE 218 238
FT DISULFID 342 369
SQ SEQUENCE 392 AA; 40113 MW; 6F699E026BF1D6A5 CRC64;

Query Match 7.3%; Score 104; DB 1; Length 392;
Best Local Similarity 22.4%; Pred. No. 0.15; 123; Indels 88; Gaps 16;
Matches 72; Conservative 38; Mismatches 123;

Qy 4 LGSRRVSKCRLVKFMSVALSVATMAVTV-----TLENTALAR-QTVSND 50
Db 32 LKHRRISKRATLAGSAVVALVAAGFTQTANASDDVPAGAKTLSADAAGKLATLDRD 91
Qy 51 VVLNDGASKYNEALAWTFNDSPNYKTLGTSQITPALFPKAGD---ILYSKLDLGRT 107
Db 92 LGADAAGSY--DATAKTL--VNVNVDGAEQVR-----QAGKARIVENSIAEL---K 139
Qy 108 TARTGLT-YANVEGSGYGRQSGFNQNPAGWTGNPNHVKYKIEWLNGLSYVGDFWNRSHL 166
Db 140 SARGTLTEKATIPGT-----SWAVDP--VSNKVLVTADSTVDGAARKKLSA 183
Qy 167 IADSLGGDALRVNAVTVGTRTQNVGGRDQ-----KGGMRY-----TQRA 205
Db 184 VVEGLGGKA-ELNRTAGEFTPLIAGGDAINGSGSRCSLGFNVKGGPYFLTAGHCTESV 242
Qy 206 QEWLEA-----NRDGYLYEVAFIYNAD-ELIPRAVNVNQSSDNTINE 248
Db 243 TSWSDTQGGSEIGANEGSSFPENDYGLVKYTSDTAHPSEVNLVDGSTQAITQAGDATVQ 302
Qy 249 KLVYNTANGYTYNHGCTPT 269
Db 303 AV-----TRSGSTTQVHDGEVT 319

RESULT 4
APU_THESA
ID APU_THESA
AC P36905;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Amylopullulanase precursor (Alpha-amylase/pullulanase) [includes:
DE Alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase);
DE Pullulanase (EC 3.2.1.41) (1,4-alpha-D-glucan glucanohydrolase)
DE (Alpha-dextrin endo-1,6-alpha-glucosidase)].

```

```

GN APU.
OS Thermoanaerobacter saccharolyticum.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacterium.
OX NCBI_TaxID=28896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B6A-RI;
RX MEDLINE=94161525; PubMed=8117096;
RA Ramesh M.V., Podkovyrov S.M., Lowe S.E., Zeikus J.G.;
RT "Cloning and sequencing of the Thermoanaerobacterium saccharolyticum
RT B6A-RI apu gene and purification and characterization of the
RT amylopullulanase from Escherichia coli.";
RL Appl. Environ. Microbiol. 60:94-101(1994).
RN [2]
RP IDENTIFICATION OF PROBABLE VECTOR CONTAMINATION.
RA Robison K.;
RL Unpublished observations (NOV-1994).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC linkages in pullulan and in amylopectin and glycogen, and the
CC alpha- and beta-limit dextrins of amylopectin and glycogen.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -1- SIMILARITY: Contains 2 fibronectin type III domains.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN BY THE PRESENCE OF
CC AN EXTRA C-TERMINAL SEGMENT OF 9 RESIDUES THAT SEEMS TO ORIGINATE
CC FROM A PUC-TYPE VECTOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L07762; AAA19800.1; ALT_SEQ.
DR HSSP; Q08751; 1BVZ.
DR InterPro; IPR005589; Alp_amyl_cat_sub.
DR InterPro; IPR006048; Alpha_amyl_C.
DR InterPro; IPR006047; Alpha_amyl_cat.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR004185; Glyco_hydro_13Ig.
DR Pfam; PF00128; alpha-amylase; 1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF02903; alpha-amylase_N; 1.
DR Pfam; PF00041; fn3; 2.
DR SMART; SM00642; Amy; 1.
DR SMART; SM00632; Amy; 1.
DR SMART; SM00060; FN3_2.
DR Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;
KW Multifunctional enzyme.
FT SIGNAL 1 35
FT CHAIN 36 1279
FT DOMAIN 929 1017
FT DOMAIN 1156 1248
FT ACT_SITE 629 629
FT ACT_SITE 658 658
FT ACT_SITE 735 735
SQ SEQUENCE 1279 AA; 142430 MW; 095CCBCA391624DD CRC64;

Query Match 7.1%; Score 100.5; DB 1; Length 1279;
Best Local Similarity 22.2%; Pred. No. 1.4;
Matches 73; Conservative 38; Mismatches 91; Indels 127; Gaps 20;

Qy 22 VALVSATMAVTVTLENTALARQTVSNDVVLNDGASKYNEALAWTFNDSPNYKTLGT 81
Db 912 IPAMSGVWLISDDGQDLTA---PQPSNVVATSGNKV---DLNSQSDGATGY-NIYR 963
Qy 82 SQITPALFPKAGD-----ILYSKLDLG-RTRTARTGLTYANVEGS 121

```

```

Db 964 SSVGGLEYEKIASNVGTTFTDNTVNGKLYVYIAISAVDELGNSEMSIDTVAYP-----A 1019
Qy 122 YGVRQSFCKNQNPAGWTGN-----PNHV-----KYIEWLNGLS----- 155
Db 1020 Y-----PIGWGNLTQVDNHNVISVNPFTEDIYAEVWADGLNSTGQGNWIAQ 1068
Qy 156 ----YVGDPMNRSHLIADSLGDLALRVNAVGTGRTQ-----NVGGRDQKG--- 197
Db 1069 LGYKYVGTVN-----DSVYGSV--YNSVGVDDSDFTWNAQYVGDIGNNDQYKASLH 1120
Qy 198 -----MRYTQRAQEWLEARNRDGYLYVEVAPIYNADELI-PRAVV-----V 237
Db 1121 LINSRSMGYLMFSDNQGSWTTTDT---LSFYVVP---SDDLKPTAPILNQGVSSRV 1174
Qy 238 SMQSSDNTINEKVLVYN-----TANGYTIN 262
Db 1175 SLTWSPSIDN--VGIIYVEIYRSDGTFN 1201

RESULT 5
ID _APU THETU STANDARD; PRT; 1861 AA.
AC P38536;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Amylopullulanase precursor (Alpha-amylase/pullulanase) (Pullulanase
DE type II) [includes: Alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan
DE glucanohydrolase); Pullulanase (EC 3.2.1.41) (1,4-alpha-D-glucan
DE glucanohydrolase); Alpha-dextrin endo-1,6-alpha-glucosidase]].
GN AMVB.
OS Thermoanaerobacter thermosulfurogenes (Clostridium
OS thermosulfurogenes).
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacterium.
OX NCBI_TaxID=33950;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 3896 / EM1;
RX MEDLINE=94252998; PubMed=8195085;
RA Matuschek M., Burchardt G., Sahm K., Bahl H.;
RT "Pullulanase of Thermoanaerobacterium thermosulfurogenes EM1
RT (Clostridium thermosulfurogenes): molecular analysis of the gene,
RT composite structure of the enzyme, and a common model for its
RT attachment to the cell surface";
RL J. Bacteriol. 176:3295-3302(1994).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic
CC linkages in oligosaccharides and polysaccharides.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic
CC linkages in pullulan and in amylopectin and glycogen, and the
CC alpha- and beta-limit dextrins of amylopectin and glycogen.
CC -!- SUBCELLULAR LOCATION: CELL-BOUND. IT C-TERMINUS MAY SERVE AS AN
CC S-LAYER ANCHOR.
CC -!- PTM: GLYCOSYLATED.
CC -!- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
CC -!- SIMILARITY: Contains 3 S-layer homology (SLH) domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57692; AAB00841.1; -
CC HSSP; Q08751; 1BVZ.
CC InterPro; IPR006589; Alp_aml_cat_sub.
CC InterPro; IPR006048; Alpha_aml1_C.
CC InterPro; IPR006047; Alpha_aml1_cat.
CC InterPro; IPR003961; FN_III.

```

```

DR InterPro; IPR004185; Glyco_hydro_13lg.
DR InterPro; IPR004193; Glyco_hydro_13N.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00128; alpha-amylase_1.
DR Pfam; PF02806; alpha-amylase_C; 1.
DR Pfam; PF02903; alpha-amylase_N; 1.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF02922; isoamylase_N; 1.
DR Pfam; PF00395; SLH; 3.
DR Pfam; SM00642; Amy; 1.
DR SMART; SM00632; Amy_C; 1.
DR SMART; SM00660; FN3; 2.
DR PROSITE; PS01072; SLH_DOMAIN; 3.
KW Hydrolase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;
KW Multifunctional enzyme; Glycoprotein.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1861 AMYLOPULLULANASE.
FT DOMAIN 928 1018 FIBRONECTIN TYPE-III 1.
FT DOMAIN 1157 1248 FIBRONECTIN TYPE-III 2.
FT ACT_SITE 628 628 BY SIMILARITY.
FT ACT_SITE 657 657 BY SIMILARITY.
FT ACT_SITE 734 734 BY SIMILARITY.
FT DOMAIN 1681 1739 SLH 1.
FT DOMAIN 1740 1803 SLH 2.
FT DOMAIN 1804 1861 SLH 3.
FT CONFLICT 1734 1734 D -> E (IN REF. 1; AAB00841).
SQ SEQUENCE 1861 AA; 206104 MW; 06C23070B453B574 CRC64;

Query Match 7.0%; Score 99; DB 1; Length 1861;
Best Local Similarity 21.4%; Pred. No. 3.1;
Matches 75; Conservative 39; Mismatches 93; Indels 144; Gaps 21;

Qy 22 VALVSATMAVTVTLENTALAROTVSDVNDVNDGASKYLNEALWTFNDSPNYKTLGT 81
Db 911 VPAMSGVMLISDDGQDLTA----PQPSNVVVVTSNGKV---DLSWLSQDAGATGYNYRS 963
Qy 82 -----SOITPALFPKAG-----DILY--SKLDELGRTRTARGTLTYANVEGSYG 123
Db 964 SVEGGLYEKIASNVTTTFEDANVTNGLKVYVIAISALDELG---NESSISNDVAVPAY- 1019
Qy 124 VROSFGKNQNPAGWTGN-----PNHV-----KYIEWLNGLS----- 155
Db 1020 -----PIGWGNLTQVDNHNHIGVDKPTEDIYAEVWADGLNSTGQGNWIAQLG 1069
Qy 156 --YVGDPMNRSHLIADSLGDLALRVNAVGTGRTQ-----NVGGRDQ----- 194
Db 1070 YKTV-----SGTVDSVYGSV--YNSVGVDDSDGFTWNAQYVGDIGNNDQYKASTPD 1121
Qy 195 KGG-----MRYTQRAQEWLEARNRDGYLYVEVAPIYNADELI-PRAVVVSVM----- 239
Db 1122 KIQWEVLMFSDNQGDWTTTST---LSFYVVP---SDDLKPTAPILNQGVTESSRSVS 1175
Qy 240 ----QSSDN-----TINEKVLVY-----NTANGVTINY 263
Db 1176 LTWNPSTDNVGIYDEIYRSDGTFNFKIATVSNEVNYIDTSVINGVTYNY 1226

RESULT 6
HLA_STAAU
ID _HLA_STAAU STANDARD; PRT; 319 AA.
AC P09616;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).
GN HLY OR HLA.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.
RC STRAIN=Wood 46;
RX MEDLINE=85053471; PubMed=6500704;

```

RA Gray G.S., Kehoe M.;  
RT "Primary sequence of the alpha-toxin gene from Staphylococcus aureus  
RL wood 46."; Infect. Immun. 46:615-618(1984).  
RN [2]  
RN REVISIONS, SEQUENCE FROM N.A.  
RP STRAIN=wood 46;  
RA Hedengrahn G.;  
RN Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.  
RP STRAIN=wood 46;  
RX MEDLINE=92268149; PubMed=1587866;  
RA Walker B., Krishnasastri M., Zorn L., Kasianowicz J., Bayley H.;  
RT "Functional expression of the alpha-hemolysin of Staphylococcus  
RP aureus in intact Escherichia coli and in cell lysates. Deletion of  
RT five C-terminal amino acids selectively impairs hemolytic  
RT activity."; J. Biol. Chem. 267:10902-10909(1992).  
RL [4]  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).  
RC STRAIN=wood 46 / ATCC 10832;  
RX MEDLINE=97102581; PubMed=8943190;  
RA Song L., Hobaugh M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;  
RT "Structure of Staphylococcal alpha-hemolysin, a heptameric  
RT transmembrane pore."; Science 274:1859-1866(1996).  
RL [5]  
RN [5]  
RP MUTAGENESIS.  
RX MEDLINE=93016135; PubMed=1400487;  
RA Walker B., Krishnasastri M., Zorn L., Bayley H.;  
RT "Assembly of the oligomeric membrane pore formed by Staphylococcal  
RT alpha-hemolysin examined by truncation mutagenesis."; J.  
RL Biol. Chem. 267:21782-21786(1992).  
RN [6]  
RP [6]  
RP MUTAGENESIS OF HISTIDINE RESIDUES.  
RX MEDLINE=94222552; PubMed=8168947;  
RA Menzies B.E., Kernodle D.S.;  
RT "Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus  
RT aureus: role of histidines in toxin activity in vitro and in a murine  
RT model."; Infect. Immun. 62:1843-1847(1994).  
RL [7]  
RN [7]  
RP MUTAGENESIS OF HISTIDINE RESIDUES.  
RX MEDLINE=94245329; PubMed=8188346;  
RA Jurach R., Hildebrand A., Hobom G., Tranum-Jensen J., Ward R.,  
RA Kehoe M., Bhakdi S.;  
RT "Histidine residues near the N-terminus of staphylococcal alpha-toxin  
RT as reporters of regions that are critical for oligomerization and  
RT pore formation."; Infect. Immun. 62:2249-2256(1994).  
RL [8]  
RN [8]  
RP MUTAGENESIS.  
RX MEDLINE=96032742; PubMed=7559447;  
RA Walker B., Bayley H.;  
RT "Key residues for membrane binding, oligomerization, and pore forming  
RT activity of staphylococcal alpha-hemolysin identified by cysteine  
RT scanning mutagenesis and targeted chemical modification."; J.  
RL Biol. Chem. 270:23055-23071(1995).  
RN [9]  
CC -1- FUNCTION: ALPHA-TOXIN BINDS TO THE MEMBRANE OF EUKARYOTIC CELLS  
CC RESULTING IN THE RELEASE OF LOW-MOLECULAR WEIGHT MOLECULES AND  
CC LEADING TO AN EVENTUAL OSMOTIC LYSIS. HEPTAMER OLIGOMERIZATION  
CC AND PORE FORMATION IS REQUIRED FOR LYTIC ACTIVITY.  
CC -1- SUBUNIT: SELF-ASSEMBLES TO FORM FIRST, A NONLYTIC OLIGOMERIC  
CC INTERMEDIATE, AND THEN, A MUSHROOM-SHAPED HOMHEPTAMER STRUCTURE  
CC OF 100 ANGSTROMS IN LENGTH AND UP TO 100 ANGSTROMS IN DIAMETER.  
CC -1- SUBCELLULAR LOCATION: SECRETED AS A MONOMER. AFTER OLIGOMERIZATION  
CC AND PORE FORMATION, THE COMPLEX IS TRANSLOCATED ACROSS THE  
CC BILAYER, PROBABLY VIA THE GLY-RICH DOMAIN OF EACH STRAND.  
CC -1- DOMAIN: THE MUSHROOM-SHAPED HEPTAMER IS COMPOSED OF A CAP DOMAIN  
CC (COMPRISING 7 BETA SANDWICHES AND THE AMINO LATCHES OF EACH  
CC PROTONER), 7 RIM REGIONS WHOSE PROTRUDING STRANDS MAY INTERACT  
CC WITH THE MEMBRANE BILAYER, AND THE STEM DOMAIN (52 ANGSTROMS IN

CC LENGTH, 26 ANGSTROMS IN DIAMETER) WHICH FORMS THE TRANSMEMBRANE  
CC PORE.  
CC -1- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; X01645; CAA25801.1; --  
CC PIR; M90536; AAA28598.1; --  
CC PDB; 7AHL; 14-JAN-98.  
CC InterPro; IPR005831; Aer hem.  
CC InterPro; IPR005830; Aer hem leuk.  
CC InterPro; IPR001340; Hemlyan\_pore.  
CC InterPro; IPR003963; Staph\_bicn\_txn.  
CC Pfam; PF01117; Aerolysin; 1.  
CC PRINTS; PR01468; BICOMPOTOXIN.  
CC TIGRFAMS; TIGR01002; hlyII; 1.  
CC PROSITE; PS00274; AEROLYSIN; 1.  
CC Hemolysis; Toxin; Signal; 3D-structure.  
CC SIGNAL 1 26  
CC CHAIN 27 319 ALPHA-HEMOLYSIN.  
CC DOMAIN 145 169 GLY-RICH.  
CC SITE 159 160 CLEAVAGE OF MONOMERS BY PROTEINASE K  
CC SITE 161 162 (MINOR SITE) CLEAVAGE OF MONOMERS BY PROTEINASE K  
CC SITE 164 165 CLEAVAGE OF MONOMERS BY PROTEINASE K  
CC SITE 165 166 CLEAVAGE OF MONOMERS BY PROTEINASE K  
CC MUTAGEN 61 61 H-SL: NO OLIGOMERIZATION NOR HEMOLYTIC  
CC MUTAGEN 61 61 ACTIVITY.  
CC MUTAGEN 61 61 H-R: NO OLIGOMERIZATION NOR HEMOLYTIC  
CC MUTAGEN 74 74 ACTIVITY.  
CC MUTAGEN 170 170 H-SL: 7% OF NORMAL HEMOLYTIC ACTIVITY.  
CC MUTAGEN 285 285 H-SL: 16% OF NORMAL HEMOLYTIC ACTIVITY.  
CC HELIX 28 31 H-SL: 46% OF NORMAL HEMOLYTIC ACTIVITY.  
CC TURN 32 32  
CC STRAND 33 33  
CC TURN 35 38  
CC STRAND 40 40  
CC TURN 42 43  
CC STRAND 47 55  
CC TURN 56 59  
CC STRAND 60 69  
CC TURN 72 73  
CC STRAND 74 74  
CC STRAND 77 87  
CC STRAND 92 92  
CC STRAND 96 96  
CC TURN 99 100  
CC STRAND 101 115  
CC TURN 118 119  
CC STRAND 123 128  
CC STRAND 135 153  
CC TURN 154 155  
CC STRAND 158 177  
CC STRAND 179 184  
CC STRAND 190 197  
CC STRAND 200 202  
CC TURN 203 204  
CC STRAND 205 208  
CC TURN 210 211  
CC STRAND 214 214  
CC TURN 215 217  
CC STRAND 218 218

```
FT TURN 221 222
FT STRAND 223 223
FT HELIX 232 234
FT STRAND 236 236
FT HELIX 239 241
FT TURN 244 247
FT STRAND 248 248
FT STRAND 250 250
FT STRAND 254 260
FT TURN 262 263
FT STRAND 268 286
FT STRAND 291 311
FT TURN 312 315
FT STRAND 316 319
SQ SEQUENCE 319 AA; 35904 MW; 6711C415DF7EBF30 CRC64;

Query Match 6.7%; Score 95.5; DB 1; Length 319;
Best Local Similarity 18.2%; Pred. No. 0.62;
Matches 54; Conservative 52; Mismatches 98; Indels 93; Gaps 12;

Qy 17 VKFSMVALVSATMAVTTTLENTALARTQVSNVDVNLNDGASKYLNEALAWTFNDSPNY 76
Db 1 MKRIVSVTTLLIGSLMNPVA-----GAA-----DSDINI 33

Qy 77 KTLGTSQITPALFPKAGD-----ILYKSLDELGRTRTARGTLTYANVEGSYG 123
Db 34 KT-GTTDLSNTTVKTDGLVTDYKENGHMKVVFSDIDDKNNKLLVIRTKGTIAGQYR 92

Qy 124 VRQSGKQNPAGWTGNPNHVKYKIEWL-NGLSVYGVDFWNRSHL-----IADSLGGD 174
Db 93 VYSEBEGANKSLAW---PSAFKVLQLPDNEVAQISDYYPNRSIDTKYMSLTLYGFNGN 149

Qy 175 ALRVNAVTTGTQNVGG-----RDQKGMR--YEQRAQ 206
Db 150 -----VTGDDTGKIGGLIGANVSIGHTLKYVQDPFKTILSPDTKKVGKVFNNMVNQ 203

Qy 207 EWLEARNRGYLYEVAPIYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTINY 263
Db 204 NWGPFYDRSW-----NPVY-GNQLPMKTRNGSKMAADNFDLPNKASSLLSSGFSFDP 254

RESULT 7
OMPL_PROPR
ID_OMPL_PROPR STANDARD; PRT; 341 AA.
AC Q52581.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Porin-like protein L precursor.
GN OMPL.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326353; PubMed=8759872;
RA Welch T.J., Bartlett D.H.;
RT "Isolation and characterization of the structural gene for OmpL, a
RT pressure-regulated porin-like protein from the deep-sea bacterium
RT Photobacterium species strain SS9.";
RL J. Bacteriol. 178:5027-5031(1996).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane
CC (potential).
CC -!- INDUCTION: IN RESPONSE TO ELEVATED HYDROSTATIC PRESSURE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; U59311; AAB50064.1; --
DR InterPro; IPR003229; OMP_2.
DR InterPro; IPR001702; Porin_Gram-ve.
DR Pfam; PF00267; Gram-ve_porins; 1.
DR ProDom; PD00808; OMP_2; 1.
KW Transmembrane; Porin; Signal; Outer membrane.
FT SIGNAL 1 21
FT CHAIN 22 341 PORIN-LIKE PROTEIN L.
FT SEQUENCE 341 AA; 36672 MW; ED3D66A9F3DEB595 CRC64;

Query Match 6.6%; Score 94; DB 1; Length 341;
Best Local Similarity 23.9%; Pred. No. 0.91;
Matches 70; Conservative 23; Mismatches 116; Indels 84; Gaps 13;

Qy 21 MVALVSATMAVTTTTL-----ENTALARQTQVSNVDVNLND--GASKYLNEALAWTFND 71
Db 5 LIALAVALAASISSVATAAEVYSDTSSLAVGGRFEARAVLADVKNKDNVTNTASSEVS DK 64

Qy 72 SPNYKTLGTSQITP-----ALFPKAGDILYKSLDELGRTRTA-----RGLTYANVE 119
Db 65 SRVRINVAGKTDTIDEDFYGVGFEKESADSDNDE---TRYAYAGVGSQYQGLVYKAD 121

Qy 120 GSYGVRSF-----GKNQNPAGWTGNPNHVKYKIEWLNGLSYVGVDFWNRSHLIADSLGGDA 175
Db 122 GSLGMLTDFTDIMAYHGNAG-----NKLAAADRNTNNLSYVGSF-----DLNGDN 167

Qy 176 LRNVAVTGTITQNVGGDQKGMR-----YEQRAQEWLEARNPD----- 214
Db 168 LTVKA-----NYYFGGSDENEGYSAAAMYAMDGLGFGAGYGDQGGSSKNGNEDTKGKQ 222

Qy 215 --GLYLYEVAPIYNADELIPRAVVVSMOSSDNTINEKVLV-----YNTANGYT 260
Db 223 AFGAISTISDFVSG-----LYQDSRNVVNDLIDESTGYEFAAAYT 266

RESULT 8
FLGE_AQUAE
ID_FLGE_AQUAE STANDARD; PRT; 482 AA.
AC O67711;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Flagellar hook protein flgE.
GN FLGE OR AQ_1859.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -!- SIMILARITY: BELONGS TO THE FLAGELLA BASAL BODY ROD PROTEINS
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; AE000760; AAC07675.1; --
DR PIR; E70460; E70460.
DR InterPro; IPR001444; Flag_bb_rod.
DR Pfam; PF00460; flg_bb_rod; 1.
```





RA Sekiguchi J., Sekowska A., Serror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takenaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpestra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler H., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto E., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*,"  
 RL Nature 390:249-256(1997).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D84432; BAA12522.1; -;  
 DR EMBL; Z99116; CAB14415.1; -;  
 DR PIR; D69957; D69957.  
 DR Subtilist; BG1686; Yqgs.  
 DR InterPro; IPR000917; Sulfatase.  
 DR Pfam; PF00884; Sulfatase; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 5 25 POTENTIAL.  
 FT TRANSMEM 42 62 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 154 174 POTENTIAL.  
 SQ SEQUENCE 638 AA; 73207 MW; 5600A4012117101A CRC64;

Query Match 6.4%; Score 91; DB 1; Length 638;  
 Best Local Similarity 23.8%; Pred. No. 3.7;  
 Matches 57; Conservative 39; Mismatches 96; Indels 48; Gaps 14;

QY 31 VTTVLENTALARTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFP 90  
 DB 247 VILVLEST---QSFLVNEKLNCEITPFLNDFIKQSYNNVYHQT-GQGTSDSEF- 300  
 QY 91 KAGDILYKSLDELGRTRTARGTLTYANVEGSGYQVRSQKQNPAGWTGNPNHYKIEW 150  
 DB 301 NVDSNLY---PLG-----RGAVFTNAGNYMAAPEILKN---SGYSAVLHANNK--- 345  
 QY 151 NLGLSYGVDFWNRSHLIADSLGDA-LRVNA--VTGTFQNVGGRDQGGNRYTEQRAQE 207  
 DB 346 -----SFWRD-LMYDSFGYDSFFDINSYVDVTENTVGWGLKDKB-----FPEQSSE- 391  
 QY 208 WLEANRDGLYEVAPIYNADELIPRAVVMQSSDNTINEKVLVYNTANGY--TINVHN 265  
 DB 392 -LMKNLP-----QPFYSRLITLTNHPFDPDLDEEDQLIDYDQNSQTLNKFYFTVRYQD 443

RESULT 11  
 KIME METJA  
 ID KIME METJA STANDARD; PRT; 312 AA.  
 AC Q59487;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Mevalonate kinase (EC 2.7.1.36) (MK).  
 GN MKV OR Mj1087.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
 CC Methanocaldococcaceae; Methanocaldococcus.  
 CX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RT *jannaschii*,"  
 RL Science 273:1058-1073(1996).  
 CC [2]  
 CC CHARACTERIZATION.  
 CC MEDLINE=99428387; PubMed=10497066;  
 RA Huang K.-X., Scott A.I., Bennett G.N.;  
 RT "Overexpression, purification, and characterization of the  
 RT thermostable mevalonate kinase from *Methanococcus jannaschii*,";  
 RL Protein Expr. Purif. 17:33-40(1999).  
 CC -!- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-  
 CC phosphomevalonate.  
 CC -!- COFACTOR: MAGNESIUM.  
 CC -!- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE  
 CC COMPETITIVE INHIBITORS. INHIBITED BY HIGH CONCENTRATION OF ATP  
 CC AS WELL.  
 CC -!- SUBUNIT: Homodimer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -!- MISCELLANEOUS: THERMOSTABLE; THE OPTIMAL TEMPERATURE FOR ACTIVITY  
 CC IS 70-75 DEGREES CELSIUS.  
 CC -!- SIMILARITY: BELONGS TO THE GHMP KINASE FAMILY. MEVALONATE KINASE  
 CC SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; U67551; AAB99088.1; -;  
 DR PIR; F64435; F64435.  
 DR PDB; 1KX; 27-MAR-02.  
 DR TIGR; MJ1087; -;  
 DR HAMAP; MF 00217; -; 1.  
 DR InterPro; IPR001174; Galkinase.  
 DR InterPro; IPR006204; GHMP kinase.  
 DR InterPro; IPR006203; GHMPKase ATP.  
 DR InterPro; IPR006205; Mv\_Galkin.  
 DR InterPro; IPR006206; Mv\_Galkinase.  
 DR Pfam; PF00288; GHMP kinases; 1.  
 DR PRINTS; PR00960; LMPPROTEIN.  
 DR PRINTS; PR00959; MEVGALKINASE.  
 DR TIGRfam; TIGR00549; mevalon kin; 1.  
 DR PROSITE; PS00627; GHMP\_KINASES\_ATP; 1.  
 KW Transferase; Kinase; ATP-binding; Magnesium; Complete proteome;  
 KW 3D-structure.  
 FT NP BIND 104 114 ATP (POTENTIAL).  
 SQ SEQUENCE 312 AA; 35177 MW; DFF9E5B869728298 CRC64;

Query Match 6.2%; Score 88.5; DB 1; Length 312;  
 Best Local Similarity 20.7%; Pred. No. 2.4;  
 Matches 56; Conservative 42; Mismatches 96; Indels 77; Gaps 13;

QY 15 RLVKFSMVALVSATMAVT-TVTLENTALARTQVNDVLDGASKYLNEALAWTFNDSP 73  
 DB 8 KVLFGHAVVGYRAISMAIDLTSTIEIKETQ-EDEILN-----LNDLNKSLGNLNEIK 63  
 QY 74 NYTKLTGTSQITPALFPKAGDILYKSLDELGRTRTARGTLTYANVEGSGYQVRSQKQNP 133  
 DB 64 N-----INENNF---GDFKYCLC-----AIKNTLDYLNIEPKTKFNISKIP 104  
 QY 134 PAGWTGNPNHYKIEWNLGSLSYGVDFWNR-----SHLIADSLGDAIRVNAVTC 183  
 DB 105 ISCGLGSSASIT-----IGTIKAVSGFYNNKELKDDDEIAKGLYMWVEIKQKA-----SITD 155

QY 184 TRTONVGG-----RDQKGMNR-----YTEORAEQWLEARNRGYLYVEVAP 223  
 Db 156 TSTITYKGLILEKNNFKRIKGFSEFELKCKFLIVYAEKRRKKTAE-----LVNEVAK 209  
 QY 224 IYNADLIPRAVWVSMQSSDNTINEKVLVYN 254  
 Db 210 IENKDEI-----FKEIDKVIDEALIKN 232

RESULT 12  
 ID GTFB STRMU STANDARD; PRT: 1476.AA.  
 AC P08987; O69381; O69384; O69387; O69390; O69396;  
 DT 01-NOV-1988 (Rel. 02, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)  
 DE (Sucrose 6-glucosyltransferase).  
 GN GTFB OR SMU.1004.  
 OS Streptococcus mutans.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=11309;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GS-5;  
 RX MEDLINE=87308013; PubMed=3040685;  
 RA Shiroza T., Ueda S., Kuramitsu H.K.;  
 RT "Sequence analysis of the gtfB gene from Streptococcus mutans.";  
 RL J. Bacteriol. 169:4263-4270(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,  
 RC MT4467 / Serotype E, and MT8148 / Serotype C;  
 RX MEDLINE=98231643; PubMed=9570124;  
 RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,  
 RA Kimura S., Hamada S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT Streptococcus mutans";  
 RL FEMS Microbiol. Lett. 161:331-336(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UA159 / ATCC 700610 / Serotype C;  
 RX MEDLINE=22295063; PubMed=12397186;  
 RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  
 RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;  
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental  
 RT pathogen.";

1- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 2- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-  
 fructose + {(1,6)-alpha-D-glucosyl} (N+1).  
 3- SUBCELLULAR LOCATION: Secreted.  
 4- DISEASE: DENTAL CARIES.  
 5- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA  
 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES  
 WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH  
 FORMS OF GLUCANS.  
 6- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
 BINDING PROTEIN FROM S. MUTANS.  
 7- SIMILARITY: Contains 10 cell wall binding repeats.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL; M17361; AAA8588.1; --  
 DR EMBL; D88651; BAA26101.1; --  
 DR EMBL; D88654; BAA26105.1; --  
 DR EMBL; D88657; BAA26109.1; --  
 DR EMBL; D88660; BAA26113.1; --  
 DR EMBL; D89977; BAA26119.1; --  
 DR EMBL; AE014940; AAN58705.1; --  
 DR InterPro; IPR002479; CW\_binding.  
 DR InterPro; IPR003318; Glyco\_hydro\_70.  
 DR Pfam; PF01473; CW\_binding\_1; 1.  
 DR Pfam; PF02324; Glyco\_hydro\_70; 1.  
 KW Transferase; Glycosyltransferase; Signal; Repeat; Dental caries;  
 KW Complete proteome.  
 FT SIGNAL 1 34 POTENTIAL.  
 FT CHAIN 35 1476 GLUCOSYLTRANSFERASE-I.  
 FT DOMAIN 35 1051 CATALYTIC (APPROXIMATE).  
 FT DOMAIN 1097 1476 GLUCAN-BINDING (APPROXIMATE).  
 FT REPEAT 1097 1130 A REPEAT.  
 FT DOMAIN 1161 1470 5 X TANDEM REPEATS.  
 FT REPEAT 1161 1210 1.  
 FT REPEAT 1225 1275 2.  
 FT REPEAT 1290 1340 3.  
 FT REPEAT 1355 1405 4.  
 FT REPEAT 1420 1470 5.  
 FT VARIANT 62 62 S -> T (IN STRAIN MT4239).  
 FT VARIANT 65 65 T -> I (IN STRAIN GS-5).  
 FT VARIANT 68 68 V -> A (IN STRAINS GS-5, MT4245, MT4251,  
 FT VARIANT 78 78 Q -> P (IN STRAIN MT4251).  
 FT VARIANT 86 86 I -> S (IN STRAINS GS-5, MT4245, MT4251,  
 FT VARIANT 89 89 MT4467 AND MT8148).  
 FT VARIANT 168 168 S -> F (IN STRAIN MT4251).  
 FT VARIANT 276 276 K -> N (IN STRAIN MT4251).  
 FT VARIANT 399 399 S -> D (IN STRAINS GS-5, MT4467 AND  
 FT VARIANT 474 474 MT8148).  
 FT VARIANT 512 512 N -> R (IN STRAIN MT4239).  
 FT VARIANT 519 512 I -> T (IN STRAIN MT4239).  
 FT VARIANT 701 519 K -> R (IN STRAIN MT8148).  
 FT VARIANT 708 701 F -> Y (IN STRAIN MT8148).  
 FT VARIANT 708 708 T -> I (IN STRAIN MT8148).  
 FT VARIANT 938 708 A -> V (IN STRAIN MT8148).  
 FT VARIANT 952 938 Q -> L (IN STRAIN MT8148).  
 FT VARIANT 963 957 FGRPE -> YGTPVA (IN STRAINS GS-5, MT4239  
 FT VARIANT 964 AND MT4467).  
 FT VARIANT 968 SV -> NT (IN STRAINS GS-5, MT4239 AND  
 FT VARIANT 970 MT4467).  
 FT VARIANT 1086 ADS -> VDG (IN STRAINS GS-5, MT4239 AND  
 FT VARIANT 1158 MT4467).  
 FT VARIANT 1163 A -> T (IN STRAIN MT4239).  
 FT VARIANT 1168 S -> N (IN STRAIN MT4239).  
 FT VARIANT 1182 H -> Y (IN STRAIN MT4251).  
 FT VARIANT 1234 E -> K (IN STRAIN MT8148).  
 FT VARIANT 1263 Y -> C (IN STRAIN MT8148).  
 FT VARIANT 1264 A -> P (IN STRAIN MT4239).  
 FT VARIANT 1272 R -> P (IN STRAIN MT8148).  
 FT VARIANT 1272 Y -> H (IN STRAINS GS-5 AND MT4467).  
 FT VARIANT 1329 S -> G (IN STRAINS GS-5, MT4239, MT4467  
 FT VARIANT 1394 AND MT8148).  
 FT VARIANT 1402 H -> Y (IN STRAINS GS-5 AND MT4467).  
 FT VARIANT 1459 Y -> H (IN STRAIN MT4467).  
 FT CONFLICT 570 570 R -> A (IN REF. 1).  
 FT CONFLICT 800 817 ADQDVAAASTAPSTGK -> LIKMFALRLARPHQMA  
 FT CONFLICT 1310 1310 (IN REF. 1).  
 FT CONFLICT 1476 1476 H -> L (IN REF. 1).  
 FT SEQUENCE 1476 AA; 165846 MW; 9C6E09F731B4CBFC CRC64;

Query Match  
Best Local Similarity 6.2%; Score 88.5; DB 1; Length 1476;  
Matches 67; Conservative 36; Mismatches 107; Indels 105; Gaps 16;

QY 47 VSDVVLNDGASKYLNALAWTFNDSPNY-----YKLTGTSQITPALFPKAGDI 95  
DB 1121 LSNGLQLRDALKNEEDGTAYVNDGRRYENGYYQFMSGVWRHFNNGEMSVGLTVIDGQV 1180  
QY 96 LYSKLDLGR-----TRTARGLTYANVEGSGVRSQFN-----QNPAGWTG- 139  
DB 1181 QY--FDEMGYQAKGFVTADGKIRYFDKQSGNMRYNRFIENEBGKWLVLGEDGAATG 1238  
QY 140 ---NPNHVKKYKLEWLSYVGD--WNRSHLIA--DSLGLDALRVN----- 179  
DB 1239 QTINGOHLRYF--ANGVQVKGFEVTDYGRISYDSDSGDQIRNRFVNAQGWYF 1295  
QY 180 ---AVTGTQTON-----VGR--DOKGGRYTE-----QRAQ-EW 208  
DB 1296 NGYAVTGARTINGOHLRYFRANGVQVKGFEVTDYGRISYDSDSGDQIRNRFVNAQGW 1355  
QY 209 LEANRDGY-----LYEVAPIYNADELIP-----RAVVSMSQSDNTINEKV-- 250  
DB 1356 FYEDNNGYAVTGARTINGOHLRYFRANGVQVKGFEVTDYGRISYDSDSGDQIRNRFVNA 1415  
QY 251 -----LVNTANGVTI 261  
DB 1416 AQGWFYEDNNGYAV 1430

## RESULT 13

POLN MANCV  
ID POLN MANCV STANDARD; PRT; 2208 AA.  
AC Q99014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein].  
OS Manchester virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses."  
RT Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
CC -!- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; X86560; CRA60262.1; --  
CC MEROPS; C24.UPW; --

DR InterPro; IPR004005; Calici coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Endoptase\_C24.

DR InterPro; IPR000605; RNA helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_P3D.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00916; 2CENDOPTASE.  
DR PRINTS; PR00918; CALICIVIRUS.  
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? ? COAT PROTEIN.  
FT NP\_BIND ? 2208 ATP (POTENTIAL).  
FT ACT\_SITE 408 415 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;

Query Match 6.2%; Score 88.5; DB 1; Length 2208;  
Best Local Similarity 19.58; Pred.No. 31; Mismatches 67; Indels 61; Gaps 8;  
Matches 38; Conservative 29;

QY 102 ELGRTRTARGLTYANVEGSGYV-----QSFGKNQNPAG---WTGNPNH 143  
DB 936 EIRAMRMAGGAYTHATIIIGRGVDEIIRTSPPRAPTRPQHYEEBGPATVFTQGGH 995  
QY 144 VKYKIEWLNG---LSVVGDFWNRSHLIAADSLGGLDALRVNAVGTGTQNVGGRDQKGM 199  
DB 996 IGVGVHNGNVITVTHVAS-----TSDEVNGSAFKITRTVGETT----- 1035  
QY 200 YTSQRAQEWLEARNRGYLYEV---APIYNABELIPRAVVSMSQSDNTINEKLVYNTA 256  
DB 1036 -----WVQGPFSQPHMQIGSGSPVFTTLLHPVFTI-----SEGTFFETPNI---TV 1079  
QY 257 NGYTVNHNGTPTOK 271  
DB 1080 NGPHVRIMNGYPTKK 1094

## RESULT 14

PAC ARTVI  
ID PAC ARTVI STANDARD; PRT; 802 AA.  
AC P31356;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Penicillin G acylase precursor (EC 3.5.1.11) (Penicillin G amidase)  
DE (Penicillin G amidohydrolase).  
GN PAC OR PA.  
OS Arthrobacter viscosus.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococccineae; Micrococccaceae; Arthrobacter.  
OX NCBI\_TaxID=1673;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=ATCC 15294;  
RX MEDLINE=94259306; PubMed=8200542;  
RA Konstantinovic M., Marjanovic N., Ljubijankic G., Glisin V.;  
RT "The penicillin amidase of Arthrobacter viscosus (ATCC 15294).";  
RL Gene 143:79-83(1994).  
CC -!- CATALYTIC ACTIVITY: Penicillin + H(2)O = a fatty acid anion + 6-  
CC aminopenicillanate.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA CHAIN AND A BETA CHAIN  
CC (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Extracellular (Potential).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 16:40:38 ; Search time 62.884 Seconds  
(without alignments)  
1112.084 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MLLGSRVFSKCRVLFKFS.....VNTANGYTYNHNGTPTOK 271

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mnci.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1418	100.0	271	16 Q9ACN6	Q9ACN6 streptococc
2	1414	99.7	271	16 Q54969	Q54969 streptococc
3	732	51.6	264	16 Q8K6X0	Q8K6X0 streptococc
4	413	29.1	332	16 Q8NZT6	Q8NZT6 streptococc
5	401	28.3	385	2 Q33735	Q33735 streptococc
6	399	27.4	327	2 Q93CC1	Q93CC1 streptococc
7	390	26.8	326	16 Q8K6B9	Q8K6B9 streptococc
8	190.5	13.4	252	2 Q8VQ29	Q8VQ29 streptococc
9	185.5	13.1	278	2 Q9JMN6	Q9JMN6 lactobacill
10	182.5	12.9	252	16 Q9A0M1	Q9A0M1 streptococc
11	173.5	12.2	266	16 Q8CM17	Q8CM17 streptococc
12	161.5	11.4	266	16 Q8P0B6	Q8P0B6 streptococc
13	161.5	11.4	268	16 Q9CJ26	Q9CJ26 streptococc
14	153.5	10.8	263	16 Q9CJ50	Q9CJ50 lactococcu
15	150	10.6	343	16 Q8E7K1	Q8E7K1 streptococc
16	139	9.8	285	16 Q8E5U4	Q8E5U4 streptococc

17	138.5	9.8	354	16 Q8E6F9	Q8E6F9 streptococc
18	138.5	9.8	354	16 Q8E0U1	Q8E0U1 streptococc
19	138	9.7	285	16 Q8E0E7	Q8E0E7 streptococc
20	131	9.2	261	16 Q8E6B6	Q8E6B6 streptococc
21	131	9.2	261	16 Q8E0P2	Q8E0P2 streptococc
22	121	8.5	293	16 Q8DT59	Q8DT59 streptococc
23	118.5	8.4	287	2 Q8VT23	Q8VT23 enterococcu
24	114.5	8.1	1118	2 Q9S0X0	Q9S0X0 clostridium
25	114.5	8.1	1118	2 Q9X721	Q9X721 clostridium
26	106	7.5	1233	16 Q8NSE4	Q8NSE4 corynebacte
27	105	7.4	24	2 Q9R4B2	Q9R4B2 streptococc
28	104.5	7.4	455	16 Q9ZL61	Q9ZL61 helicobacte
29	104.5	7.4	764	16 Q8XTM0	Q8XTM0 raistonia s
30	104	7.3	21	2 Q9RSX0	Q9RSX0 streptococc
31	103.5	7.3	705	16 Q8XTL2	Q8XTL2 raistonia s
32	102.5	7.2	353	10 Q9LR37	Q9LR37 arabidopsis
33	102	7.2	299	16 Q8D6V8	Q8D6V8 vibrio vuln
34	101	7.1	372	16 Q8G2H7	Q8G2H7 brucella su
35	100.5	7.1	1296	2 Q34108	Q34108 helicobacte
36	100	7.1	160	16 Q9A0I2	Q9A0I2 streptococc
37	100	7.1	928	10 Q9ZQL7	Q9ZQL7 arabidopsis
38	99.5	7.0	1291	2 Q9ZHU7	Q9ZHU7 helicobacte
39	99.5	7.0	1291	2 Q9ZHU6	Q9ZHU6 helicobacte
40	99	7.0	1364	2 Q99QC6	Q99QC6 shigella fl
41	99	7.0	1364	2 Q8VSL2	Q8VSL2 shigella fl
42	98.5	6.9	347	16 Q97FU6	Q97FU6 clostridium
43	98.5	6.9	1291	2 Q9ZHT2	Q9ZHT2 helicobacte
44	98.5	6.9	1291	2 Q9ZHU1	Q9ZHU1 helicobacte
45	98.5	6.9	1291	2 Q9LBC1	Q9LBC1 helicobacte

#### ALIGNMENTS

#### RESULT 1

Q9ACN6	PRELIMINARY;	PRT;	271 AA.
ID	Q9ACN6		
AC	Q9ACN6;		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)		
DE	Streptodornase B precursor (Mitogenic factor 25K).		
GN	DNASEB OR MF OR SPYM3 1745.		
OS	Streptococcus pyogenes, and		
OS	Streptococcus pyogenes (serotype M3).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1314, 198466;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=S.pyogenes; STRAIN=C203S;		
RX	MEDLINE=21233096; PubMed=11335140;		
RA	Gerlach D., Schmidt K.H., Fleischer B.;		
RT	"Basic streptococcal superantigens (SPEX/SMEZ or SPEC) are responsible for the mitogenic activity of the so called mitogenic factor MF.";		
RL	FEMS Immunol. Med. Microbiol. 30:209-216(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=S.pyogenes; STRAIN=MGAS315 / Serotype M3;		
RX	MEDLINE=22133808; PubMed=1212206;		
RA	Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,		
RA	Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.P., Parkins L.D.,		
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,		
RA	Schlievert P.M., Musser J.M.;		
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).		
DR	EMBL; AJ295272; CAC35734.1; --		
DR	EMBL; AF014170; AAM80352.1; --		
KW	Signal; Complete proteome.		
FT	SIGNAL	1 42	POTENTIAL.
FT	CHAIN	43 271	STREPTODORNASE B.





```
Db 5 RLRKEISIFLILLISSPISYQATLADT---KEVNANRY---NSSDTYLPALSWT 58
Qy 69 FNDSPNYKTLGTSQITPALPPKAGDILYKSLDELGRTRTARGLTYANVEGSGVRSQSF 128
Db 59 LETSPNYKVLGSGIVENLPPPKQIYVGGDLSLGRITLVRGITLTNNVLGSGYNIRKDF 118
Qy 129 --GKNQNPAGWTGNPNHVYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNAVTTGRT 186
Db 119 KRKAETLSGLGNKGNGEVYVYKIGLGDSDSYQGYFWNKSHLIADSLGGDALRVNAVTTGRT 178
Qy 187 QNVGGRQDGGMRVTEQRAQEWLEARNRQDGLYVYVAPINADSLIYANVVMSSDNTI 246
Db 179 QNVGGRSGNGMRVTEKSKWLBAHRDGLYVYVAPINADSLIYANVVMSSDNTI 238
Qy 247 NEKVLVNTANGTYTINYHNGT 267
Db 239 NEKVLVNVANGTYTIDYNOGT 259

RESULT 4
ID Q8NZT6 PRELIMINARY; PRT; 332 AA.
AC Q8NZT6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE Streptococcus.
GN SPYM18.1746.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=1917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Bares S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AE010084; AAL98274.1; -.
KW Complete proteome.
SQ SEQUENCE 332 AA; 37652 MW; 58DC9F930865FBFC0 CRC64;

Query Match 29.1%; Score 413; DB 16; Length 332;
Best Local Similarity 32.8%; Pred. No. 2.8e-24;
Matches 97; Conservative 42; Mismatches 95; Indels 62; Gaps 6;

Qy 17 VKFSWVALVSATMAVTV--TLENTALRQTVSNVDVNDGASKYLNEALWTFNDSPN 74
Db 10 IRYFTTALATLFLGLFPISKTTNSIIYAETDISN-----VMNIVQHPN 53
Qy 75 YYKTLGTSQITPALFP-----KAG-----DILYSKLDLGRTRTARGLT 114
Db 54 YYIVEGKSHLNKEFPQIYHTEKVKYKSGQSTKRVTVSDIQVSLDGYGSGEAYVIT 113
Qy 115 YANVEGSGVRSQFGKNQNPAGWTG-----NP-----NHVKYKIEW 150
Db 114 KDMIDMSAGYREKWKSPKPSGWSYFFKNTNQRAATESDYKHSKPNVSKISNNIKASIP 173
Qy 151 LNLGSLYVGDFFWNRSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEW 210
Db 174 SNGRTRGYLFDNRSHLIADSLGGRPRNNLTGRTQNVGNDRKGMQYIENKVLDDH 233
Qy 211 ANRDGILYVYVAPINADSLIYANVVMSSDNTINEKVLVNTANGTYTINYHNG 266
Db 234 KNPKVHYVYKATPVYQGSSELLPRVLVLSALSSDGTIDETVRVFNNAVAGFNIDYQNG 289
```

```
RESULT 5
ID O33735 PRELIMINARY; PRT; 385 AA.
AC O33735;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE Streptococcus.
GN SDA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RA Podbielski A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RX MEDLINE=97101060; PubMed=8945587;
RA Podbielski A., Zarges I., Floedorff A., Weber-Heymann J.;
RT "Molecular characterization of a major serotype M49 group A
RT streptococcal DNase gene (sdad).";
RL Infect. Immun. 64:5349-5356 (1996).
DR EMBL; X84793; CAA59264.1; -.
SQ SEQUENCE 385 AA; 43476 MW; 14922B0AFC3BD6D5 CRC64;

Query Match 28.3%; Score 401; DB 2; Length 385;
Best Local Similarity 32.2%; Pred. No. 3e-23;
Matches 96; Conservative 47; Mismatches 91; Indels 64; Gaps 7;

Qy 16 LVKFSWVALVSATMAVTVTLENTA---LARQTVSNVDVNDGASKYLNEALWTFNDS 72
Db 9 IHSALTALTLFLGLPIT-ENTSSIIYAESATISNN-----WSIEQH 51
Qy 73 PNYKTLGTSQI-----TPALFPKAG-----DILYSKLDLGRTRTARGLT 112
Db 52 PNYHVEGKQAQLDIKNPPELYRTTERRYKSGQSTKRVTVSDIYSLDGYGSGEAYGI 111
Qy 113 LTVANVEGSGVRSQFGKNQNPAGWTG-----NP-----NHVKYK 148
Db 112 ITKDMIDMSAGYREKWKSPKPSGWSYFFKNTNQRAATESDYKHSKPNVSKISNNIKASI 171
Qy 149 EWLNLGSLYVGDFFWNRSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEW 208
Db 172 LLSNGNVNRNGYLFDRSHLIADSLGGRPRNNLTGRTQNVGNDRKGMQYIENKVLDDH 231
Qy 209 LEARNRQDGLYVYVAPINADSLIYANVVMSSDNTINEKVLVNTANGTYTINYHNG 266
Db 232 IKENPKVHYVYKATPVYQGSSELLPRVLVLSALSSDGTIDETVRVFNNAVAGFNIDYQNG 289

RESULT 6
ID Q93CCI PRELIMINARY; PRT; 327 AA.
AC Q93CCI;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Sdalpha deoxyribonuclease.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Hasegawa T., Torii K., Hashikawa S., Iinuma Y., Ohta M.;
RT "Cloning and characterization of the deoxyribonuclease sdaIpha gene
RT from Streptococcus pyogenes.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410852; AAL02116.1; -.
```

Qy



RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RT "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*";  
RL Proc. Natl Acad. Sci. U.S.A. 98:4658-4663(2001).  
DR EMBL: AE006579; AAK34241.1; -.  
KW Complete proteome.  
SQ SEQUENCE 268 AA; 30398 MW; 24EA4762A5DDAA5E CRC64;

Query Match 11.4%; Score 161.5; DB 16; Length 268;  
Best Local Similarity 25.6%; Pred. No. 9.4e-05;  
Matches 60; Conservative 36; Mismatches 83; Indels 55; Gaps 12;

QY 21 MVALVSATMAVTTVLTENTALARQTVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
DB 15 VVILIAITFTTSTVTAARKIRNFPDPTTEILLGTKATE----- 53

QY 81 TSQITPALPKAG--DILYSKLDLGRTRTARGTLTYANV---EGSYGVRSFGKQNP 134  
DB 54 ----TPGIILPFTGSYQLVGLDLDNLRP-----TFAHQLKDDQEPNKKR-GLKFNP 101

QY 135 AGWTGNPNHVKYKIEWLNGLSYVGDFW--NRSHLIADSLGG--DALRVNAVTTGTRTQNVG 190  
DB 102 PCGW----H-NYKLTIDANGKT---TWLMDRGHLVGYQFSGLDNPK-NLVTTMTKVLNTG 150

QY 191 GRDQKG-GNRYTEQRAQEWLEARNDCYLYYEVAPIYNADELIPRAVVSMQSSD 243  
DB 151 FSDKNPLGLMYENRLDLSWALHPNFWLDYKVTVPYHKNELVPRQVLYGVGID 204

RESULT 14  
Q9CJ50 PRELIMINARY; PRT; 263 AA.

ID Q9CJ50  
AC Q9CJ50;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
DE Hypothetical protein ybE8.  
GN YBFB OR LL0156.  
OS *Lactococcus lactis* (subsp. *lactis*) (*Streptococcus lactis*).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OC NCBI\_TaxID=1360;  
RN RN  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RX MEDLINE=21235186; PubMed=11337471;  
RY Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,  
RT Weissenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium *Lactococcus*  
RT *lactis* ssp. *lactis* IL1403";  
RL Genome Res. 11:731-753(2001).  
RL EMBL: AE006253; AAK04254.1; -.  
DR Hypothetical protein; Complete proteome.  
KW SEQUENCE 263 AA; 28834 MW; EAA95B6FA317207C CRC64;

Query Match 10.8%; Score 153.5; DB 16; Length 263;  
Best Local Similarity 24.4%; Pred. No. 0.00039;  
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;

QY 68 TPDSNPYKTLGTS-QITPALFPKAGDILYSKLDLGR-----RTARGTL 113  
DB 39 TDNSQVSTKSLASVVKQAPLTFKNQKQVMWANTDALGRAVDSHIQLKDSQBPVKVREPL 98

QY 114 TYANVEGSYGVRSFGKQNPAGMTGNPNHVKYKIEWLNGLSYVGDFW--NRSHLIA--- 168  
DB 99 TY-----NPGW-----H-NYNFYKSDSGISGWMLMARGHLVGYQF 135

QY 169 DSLGDALRV-----NAVGTGRTQNVGRDQKGMRYTEQRAQEWLEARNRQDLYY 219  
DB 136 SGLNNEARNLPETAWFNGGNFTGNDGNT-----ASMLYYENRLDLSWALHPNYLDY 189

QY 220 EVAPLYNADELIPRAV---VVSMSQSSDNTINPK-----VLYVNTANGVTT 261

```

Db 190 QVTPLYEGNELLPRQIRLAYVGDKNQGTLSIKLGGREKSGNGGATVVVLDNVAPNAKI 249
Qy 262 NYHNGT 267
Db 250 NYADGT 255

RESULT 15
Q8E7K1 PRELIMINARY; PRT; 343 AA.
AC Q8E7K1
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN G8S0153.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NM316 / Serotype III;
RX MEDLINE=2242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lalloui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RL invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766844; CAD45798.1; -.
DR SgaliList; gbs0153; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 343 AA; 38877 MW; 11BF35A12D959DF3 CRC64;

Query Match 10.6%; Score 150; DB 16; Length 343;
Best Local Similarity 25.5%; Pred.No. 0.001;
Matches 62; Conservative 33; Mismatches 102; Indels 46; Gaps 11;

Qy 51 VVLDGASKYLNALAWTFNDSPN---YYKTLGTSQITPALFPKAGDILYSKLDLGRTR 107
Db 16 VVLCVGLQQNVAKAKHYKTTSHVETQYVSTSSKKILP--FTHNKQIKVGPLDNLGRA- 72
Qy 108 TARGTLTYANVEGSGYVRQSPGKNQ---NPAGWTGNPNHVKY-----KIEWLNGLSYVGD 159
Db 73 -----TYSHIQLRDADDEPKIKRRLTYNPTGW----HNYKFTTEKGTTLW----- 114
Qy 160 FWNRSHLIADSLGG-DALRVNAVGTGTRTQNVG-GRDQKGMRYTEQRAQEWLEANRDGYL 217
Db 115 -MDRGHLVGYQFSGMNVNPNLVTMTKYLNTGFSENNPDGMWLYYENRLDSWLANHKNFWL 173
Qy 218 YVEVAPIYNADELIPRAVVSMQSDNTINEKVLVYNTANG-----YTINYHNGTP 268
Db 174 DYKVTPIYEGNLPVSRVELQYVGIDK--QGKLLLEIKLGGKQETDEYGVTTVTLENTSP 231
Qy 269 TQK 271
Db 232 LAK 234
```

Search completed: January 5, 2004, 18:41:34  
Job time : 67.884 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 5, 2004, 15:08:48 ; Search time 59.0424 Seconds  
(without alignments)  
615.632 Million cell updates/sec

Title: US-08-482-785-9

Perfect score: 1220

Sequence: 1 RQTQVNDVLDGASKYLN.....VYNTANGYTYNYNGPTQK 229

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

```
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1220	100.0	229	16	AA70701
2	1220	100.0	293	16	AA70702
3	1216	99.7	271	23	ABP30017
4	1214	99.5	271	17	ABP30017
5	1195	98.0	271	15	ABP30017
6	258	21.1	91	16	AA70700
7	253	20.7	90	17	AA70700
8	202	16.6	38	17	AA70700
9	171.5	14.1	252	23	ABP27393
					Recombinant DNA-ase
					DNA-ase-B. Streptococcus
					Streptococcus poly
					S. pyogenes DNaseB
					Mitogenic factor a
					DNA-ase-B2 N-termi
					S. pyogenes DNase
					Recombinant DNaseB
					Streptococcus poly

10	158.5	13.0	268	23	ABP25596
11	153.5	12.6	263	23	ABP25596
12	134.5	11.0	262	21	ABP25596
13	134	11.0	354	23	ABP27392
14	131	10.7	157	23	ABP26062
15	131	10.7	261	23	ABP30074
16	127	10.4	271	23	ABP27570
17	127	10.4	285	23	ABP29813
18	126.5	10.4	247	20	AA722337
19	126.5	10.4	247	22	AA722337
20	124.5	10.2	242	20	AA722337
21	124.5	10.2	242	22	AA722337
22	122.5	10.0	274	21	AA781780
23	122.5	10.0	274	21	AA781780
24	106	8.7	23	17	AA788824
25	106	8.7	1233	22	AA788824
26	104.5	8.6	455	19	AA710950
27	104.5	8.6	455	20	AA717182
28	104.5	8.6	486	18	AA720586
29	100	8.2	160	23	ABP26063
30	94	7.7	293	20	AA701365
31	92	7.5	867	23	AA701365
32	91.5	7.5	1118	22	AA765579
33	91	7.5	777	17	AA781810
34	91	7.5	1429	21	AA793606
35	88.5	7.3	1008	19	AA763721
36	88	7.2	2137	23	ABP39618
37	86	7.0	1136	22	AA76854
38	86	7.0	1136	22	AA76854
39	86	7.0	1179	22	AA76854
40	86	7.0	1182	24	ABP18970
41	85.5	7.0	380	23	AA79286
42	85.5	7.0	1185	23	ABP40431
43	85.5	7.0	1475	23	AA798027
44	85.5	7.0	1475	23	AA798027
45	85.5	7.0	1475	23	AA798030
					AA798031

## ALIGNMENTS

### RESULT 1

AA70701

ID AA70701 standard; protein; 229 AA.

XX AA70701;

XX AA70701;

XX AA70701;

DT 25-MAR-2003 (updated)

DT 14-AUG-1995 (first entry)

XX AA70701;

DE Recombinant DNA-ase-B.

XX DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.

XX DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.

XX Streptococcus pyogenes.

OS Streptococcus pyogenes.

PN WO9500650-A1.

XX WO9500650-A1.

PD 05-JAN-1995.

XX 18-MAY-1994; 94WO-US05626.

XX 23-JUN-1993; 93US-0082845.

XX (BECI ) BECKMAN INSTR INC.

XX Adams CW, Belai CM, Pang PPY;

XX WPI; 1995-052087/07.

XX New DNA encoding Streptococcus pyogenes DNase B - for diagnosing

PT S. pyogenes infection, also new promoter for expressing other

PT proteins

XX PS Claim 1; Fig 4; 97pp; English.  
XX CC This is the full-length sequence of Streptococcus pyogenes  
CC DNA-ase-B, which is a marker of S. pyogenes infection. The protein  
CC may be used as a diagnostic agent or vaccine for S. pyogenes, or  
CC may be used as an aerosol to treat excessive lung viscosity, e.g.  
CC in cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX SQ Sequence 229 AA;  
Query Match 100.0%; Score 1220; DB 16; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.7e-120;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 60  
DB 1 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 60  
QY 61 LGRTTRTARGTLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKKYKIEWNLGLSYVGDFWN 120  
DB 61 LGRTTRTARGTLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKKYKIEWNLGLSYVGDFWN 120  
QY 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
DB 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
QY 181 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229  
DB 181 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229

RESULT 2  
AAR70702  
ID AAR70702 standard; Protein; 293 AA.  
XX AC AAR70702;  
XX DT 25-MAR-2003 (updated)  
XX DT 15-AUG-1995 (first entry)  
XX DE DNA-ase-B.  
XX KW DNA-ase-B; diagnostic; vaccine; cystic fibrosis therapy.  
XX OS Streptococcus pyogenes (ATCC 14289).  
XX EH Key Location/Qualifiers  
XX FT Protein 46..293  
XX FT Peptide 1..45 /note= "mature protein"  
XX FT misc\_difference 274 /note= "leader peptide: claim 11"  
XX FT /note= "in-frame stop codon"  
XX PN WO9500650-A1.  
XX XX  
XX PD 05-JAN-1995.  
XX PF 18-MAY-1994; 94WO-US05626.  
XX PR 23-JUN-1993; 93US-0082845.  
XX PA (BECI ) BECKMAN INSTR INC.  
XX PI Adams CW, Belei CW, Pang PPY;  
XX DR WPI; 1995-052087/07.  
XX DR N-PSDB; AAQ85037.  
XX PT New DNA encoding Streptococcus pyogenes DNase B - for diagnosing  
XX S. pyogenes infection, also new promoter for expressing other

PT proteins  
XX XX Disclosure; Fig 5; 97pp; English.  
XX CC The sequence shows a gene product corresponding to a Streptococcus  
CC pyogenes DNA-ase-B. The protein is useful as a diagnostic agent,  
CC vaccine or as an aerosol to treat excessive lung viscosity, e.g. in  
CC cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX SQ Sequence 293 AA;  
Query Match 100.0%; Score 1220; DB 16; Length 293;  
Best Local Similarity 100.0%; Pred. No. 2.5e-120;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 60  
DB 45 RQTQVNDVLDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 104  
QY 61 LGRTTRTARGTLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKKYKIEWNLGLSYVGDFWN 120  
DB 105 LGRTTRTARGTLTYANVEGSYGVRSFGKQNPAGWTGNPNHVKKYKIEWNLGLSYVGDFWN 164  
QY 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
DB 165 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 224  
QY 181 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229  
DB 225 PIYNADLIPRAVVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 273

RESULT 3  
ABP30017  
ID ABP30017 standard; Protein; 271 AA.  
XX AC ABP30017;  
XX DT 02-JUL-2002 (first entry)  
XX DE Streptococcus polypeptide SEQ ID NO 9210.  
XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
XX KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
XX KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
XX OS Streptococcus pyogenes.  
XX PN WO200234771-A2.  
XX PD 02-MAY-2002.  
XX PF 29-OCT-2001; 2001WO-GB04789.  
XX PR 27-OCT-2000; 2000GB-0026333.  
XX PR 24-NOV-2000; 2000GB-0028727.  
XX PR 07-MAR-2001; 2001GB-0005640.  
XX PA (CHIR-) CHIRON SPA.  
XX PA (GENO-) INST GENOMIC RES.  
XX PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;  
XX PI Tettelin H;  
XX DR WPI; 2002-352536/38.  
XX DR N-PSDB; ABN70648.  
XX PT New Streptococcus protein for the treatment or prevention of infection  
XX or disease caused by Streptococcus bacteria, such as meningitis, and  
XX for detecting a compound that binds to the protein -  
XX Claim 1; Page 4045; 4525pp; English.  
PS



XX The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A Streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 XX  
 SQ Sequence 271 AA;

Query Match 99.7%; Score 1216; DB 23; Length 271;  
 Best Local Similarity 99.6%; Pred. No. 5.8e-120;  
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROTQSVNDVNLGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLDE 60  
 Db 43 ROTQSVNDVNLGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLDE 102  
 Qy 61 LGTRTRARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 120  
 Db 103 LGTRTRARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 162  
 Qy 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGVLYYEVA 180  
 Db 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGVLYYEVA 222  
 Qy 181 PIYNADLIPRAVVSQSSDNTINEKLVYNTANGTYTINYHNGTPTQK 229  
 Db 223 PIYNADLIPRAVVSQSSDNTINEKLVYNTANGTYTINYHNGTPTQK 271

RESULT 4  
 AAR88823  
 ID AAR88823 standard; Protein; 271 AA.  
 XX  
 AC AAR88823;  
 DT 25-JUN-1996 (first entry)  
 XX  
 DE S. pyogenes DNaseB and leader sequence.  
 XX  
 KW DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
 KW diagnosis; Escherichia coli.  
 XX  
 OS Streptococcus pyogenes strain ATCC 14289.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 1..43  
 FT Protein /label= Sig\_peptide  
 FT 44..271  
 FT /label= Mat\_protein  
 XX  
 PN WO9606174-A1.  
 XX  
 PD 29-FEB-1996.  
 XX  
 PF 18-AUG-1994; 94WO-US09450.  
 XX  
 PR 18-AUG-1994; 94WO-US09450.  
 XX  
 PA (BECI ) BECKMAN INSTR INC.  
 XX

PI Adams CW, Belei MC, Pang PPY;  
 XX  
 DR WPI; 1996-151377/15.  
 DR N-PSDB; AAT12774.  
 XX  
 PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
 PT prodn. of the enzyme in other bacteria, useful in immunoassays or  
 PT for treating cystic fibrosis  
 XX  
 PS Claim 1; Page 67-70; 115pp; English.  
 XX  
 CC Streptococcus pyogenes DNase B, including the leader peptide,  
 CC has the amino acid sequence given in AAR88821. The enzyme can  
 CC be obtd. on a large scale by expression of encoding DNA (AAT12774)  
 CC in transformed host cells, esp. Escherichia coli. Inclusion  
 CC of the leader peptide facilitates purification of the recombinant  
 CC enzyme. The DNase B is useful in immunoassays to detect  
 CC anti-DNase B antibodies in serum as a marker for S. pyogenes  
 CC infection, and is also useful as a vaccine or for treatment, via  
 CC aerosol delivery, of cystic fibrosis.  
 XX  
 SQ Sequence 271 AA;

Query Match 99.5%; Score 1214; DB 17; Length 271;  
 Best Local Similarity 99.8%; Pred. No. 9.4e-120;  
 Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ROTQSVNDVNLGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLDE 60  
 Db 43 ROTQSVNDVNLGASKYLNEALAWTFNDSPNYKTLGTSQITPALPPKAGDILYSKLDE 102  
 Qy 61 LGTRTRARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 120  
 Db 103 LGTRTRARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 162  
 Qy 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGVLYYEVA 180  
 Db 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGVLYYEVA 222  
 Qy 181 PIYNADLIPRAVVSQSSDNTINEKLVYNTANGTYTINYHNGTPTQK 229  
 Db 223 PIYNADLIPRAVVSQSSDNTINEKLVYNTANGTYTINYHNGTPTQK 271

RESULT 5  
 AAR58702  
 ID AAR58702 standard; Protein; 271 AA.  
 XX  
 AC AAR58702;  
 DT 25-MAR-2003 (updated)  
 DT 29-MAR-1995 (first entry)  
 XX  
 DE Mitogenic factor associated with group A Streptococci.  
 XX  
 KW mitogenic factor; microdetection; group A streptococci; spe;  
 KW erythrogenic toxin; streptococcal pyrogenic exotoxin; blastogens;  
 KW scarlet fever toxin; erythematous skin reaction; infectious disease;  
 KW delayed hypersensitivity; ss.  
 XX  
 OS Streptococcus pyogenes.  
 XX  
 PN EP613947-A2.  
 XX  
 PD 07-SEP-1994.  
 XX  
 PF 31-JAN-1994; 94EP-0101386.  
 XX  
 PR 01-FEB-1993; 93JP-0037383.  
 XX  
 PA (SHIO ) SHIONOGI & CO LTD.  
 XX  
 PI Hara A, Hinuma Y, Igarashi H, Iwasaki M, Kishishita M;

PI Okumura K, Takeda Y, Yutsudo T;  
XX WPI; 1994-272994/34.  
DR N-PSDB; AAQ71612.  
XX  
PT New mitogenic factor gene from Streptococcus pyogenes - used to  
PT develop prods. for the early diagnosis of infectious disease  
PT caused by gp A streptococci  
XX  
XX Claim 7; Page 12-13; 20pp; English.  
PS  
XX AAR58702 shows a mitogenic factor which exhibits rabbit peripheral  
CC blood lymphocyte mitogenicity and/or DNA hydrolysing activity. It is  
CC strongly associated with group A Streptococci and the nucleotide  
CC sequences can be used for the microdetection of the gene and provide  
CC an early diagnosis of infectious disease caused by the bacteria.  
CC (See also AAQ71613-26).  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 271 AA;  
Query Match 98.0%; Score 1195; DB 15; Length 271;  
Best Local Similarity 98.3%; Pred. No. 9.6e-118;  
Matches 225; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALPKAGDILYKLDE 60  
Db 43 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALPKAGDILYKLDE 102  
QY 61 LGRTETAGTLTYANVEGSGYGVRSFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDFWN 120  
Db 103 LGRTETAGTLTYANVEGSGYGVRSFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDFWN 162  
QY 121 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEORAEWLEANDRDGILYVEVA 180  
Db 163 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEORAEWLEANDRDGILYVEVA 222  
QY 181 PTYNADLIPRAVVVSMOSSDNTINEKVLVNTANGYTIYVHNGTPTOK 229  
Db 223 PTYNADLIPRAVVVSMOSSDNTINEKVLVNTANGYTIYVHNGTPTOK 271  
RESULT 6  
AAR70700  
ID AAR70700 standard; Protein; 91 AA.  
AC  
XX  
XX AAR70700;  
XX  
DT 25-MAR-2003 (updated)  
DT 14-AUG-1995 (first entry)  
XX  
XX DNA-ase-B2 N-terminal fragment.  
XX  
XX DNA-ase-B2 N-terminal fragment; diagnostic; vaccine; cystic  
KW fibrosis therapy.  
KW  
XX Streptococcus pyogenes.  
OS  
XX  
XX WO9500650-A1.  
PN  
XX  
XX 05-JAN-1995.  
PD  
XX  
XX 18-MAY-1994; 94WO-US05626.  
PF  
XX  
XX 23-JUN-1993; 93US-0082845.  
PR  
XX  
XX (BECI ) BECKMAN INSTR INC.  
PA  
XX  
XX Adams CW, Belei CM, Pang PPY;  
PI  
XX WPI; 1995-052087/07.  
DR  
XX N-PSDB; AAQ85036.  
XX

PT New DNA encoding Streptococcus pyogenes DNase B - for diagnosing  
PT S. pyogenes infection, also new promoter for expressing other  
PT proteins  
XX  
XX Claim 3; Fig 3; 97pp; English.  
PS  
XX The sequence represents the N-terminal fragment of Streptococcus  
CC pyogenes DNA-ase-B, which is a marker of S. pyogenes infection.  
CC The protein may be used as a diagnostic agent or vaccine for S.  
CC pyogenes, or may be used as an aerosol to treat excessive lung  
CC viscosity, e.g. in cystic fibrosis.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 91 AA;  
Query Match 21.1%; Score 258; DB 16; Length 91;  
Best Local Similarity 100.0%; Pred. No. 2.5e-19;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALFPK 49  
Db 43 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALFPK 91  
RESULT 7  
AAR88821  
ID AAR88821 standard; Protein; 90 AA.  
XX  
XX AAR88821;  
AC  
XX 25-JUN-1996 (first entry)  
DT  
XX  
XX S. pyogenes DNase B partial sequence.  
DE  
XX  
XX DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;  
KW diagnosis; Escherichia coli.  
KW  
XX Streptococcus pyogenes strain ATCC 14289.  
OS  
XX  
XX  
XX Key Location/Qualifiers  
FT Peptide 1..43  
FT /label= Sig\_peptide  
XX  
XX WO9606174-A1.  
PN  
XX  
XX 29-FEB-1996.  
PD  
XX  
XX 18-AUG-1994; 94WO-US09450.  
PF  
XX  
XX 18-AUG-1994; 94WO-US09450.  
PR  
XX  
XX (BECI ) BECKMAN INSTR INC.  
PA  
XX  
XX Adams CW, Belei MC, Pang PPY;  
PI  
XX  
XX WPI; 1996-151377/15.  
DR  
XX N-PSDB; AAT12773.  
DR  
XX  
XX New DNA encoding Streptococcus pyogenes DNase B - for recombinant  
PT prodn. of the enzyme in other bacteria, useful in immunoassays or  
PT for treating cystic fibrosis  
PT  
XX  
XX Example 2; Fig 3; 115pp; English.  
PS  
XX  
XX A partial sequence (AAR88821) for Streptococcus pyogenes DNase B  
CC is the product of DNA clone lambda 2-6 (AAT12773) isolated from a  
CC S. pyogenes DNA library. A full-length sequence is given in  
CC AAR88823. The insert of the DNase B-encoding DNA clone can be  
CC transferred to expression plasmids for large-scale recombinant  
CC DNase prodn. in E. coli or other bacterial hosts. It is used in  
CC immunoassays to detect anti-DNase B antibodies in serum as a  
CC marker of S. pyogenes infection, and is also useful as a vaccine  
CC or for cystic fibrosis treatment via aerosol delivery.  
CC

```

XX Sequence 90 AA;
SQ
Query Match 20.7%; Score 253; DB 17; Length 90;
Best Local Similarity 100.0%; Pred. No. 8.3e-19;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALFP 48
DB 43 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKYTLGTSQITPALFP 90

RESULT 8
AAR88825
ID AAR88825 standard; Peptide; 38 AA.
XX
AC AAR88825;
XX
DT 25-JUN-1996 (first entry)
XX
DE Recombinant DNaseB N-terminal peptide.
XX
KW DNase B; nuclease; cystic fibrosis; vaccine; immunoassay;
KW diagnosis; Streptococcus pyogenes.
XX
OS Synthetic.
XX
PN WO9606174-A1.
XX
PD 29-FEB-1996.
XX
PF 18-AUG-1994; 94WO-US09450.
XX
PR 18-AUG-1994; 94WO-US09450.
XX
PA (BECI ) BECKMAN INSTR INC.
XX
PI Adams CW, Belei MC, Pang PPY;
XX
DR WPI; 1996-151377/15.
XX
PT New DNA encoding Streptococcus pyogenes DNase B - for recombinant
PT prodn. of the enzyme in other bacteria, useful in immunoassays or
PT for treating cystic fibrosis
XX
PS Example 6; Page 48; 115pp; English.
XX
CC The N-terminal sequence (AAR88825) of a Streptococcus pyogenes
CC recombinant DNase B expressed in Escherichia coli transformants was
CC detd. A modified DNase B gene (AAT12780) was produced by PCR
CC amplification of the native DNase B gene (AAT12774) in clone lambda 2-6.
CC The gene was inserted into vector del-33 for expression in E. coli.
CC Modification of the gene resulted in the presence of an additional
CC Arg residue at the DNase B N-terminus (see also AAR88824).
XX
SQ Sequence 38 AA;
Query Match 16.6%; Score 202; DB 17; Length 38;
Best Local Similarity 100.0%; Pred. No. 5.6e-14;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKYTLG 38
DB 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKYTLG 38

RESULT 9
ABP27393
ID ABP27393 standard; Protein; 252 AA.
XX
AC ABP27393;
XX
DT 02-JUL-2002 (first entry)
XX

```

---

```

XX Streptococcus polypeptide SEQ ID NO 3962.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Frazer C;
PI Tetelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN68024.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3550; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 252 AA;
Query Match 14.1%; Score 171.5; DB 23; Length 252;
Best Local Similarity 27.6%; Pred. No. 1.6e-09;
Matches 62; Conservative 37; Mismatches 81; Indels 45; Gaps 13;

QY 33 YKTLGTQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEGSGVRSFGK-NQN 91
DB 44 HYKNTVSSKLLP--FTANYQLQLGELNLN-----RATFSHIQLQDRHETKQVTRKINYD 96
QY 92 PAGWTGNPNHVYKIEWLNG--LSVVGDFWNRSHLIADSLGG--DALRVNAVGTGRTQNV 147
DB 97 PVGW-----H-NYQPPYGDGSKSWV---MNRHVLVGYQFCGLNDEPR-NLVAMTAWLT 146
QY 148 G---GRDQKG--GMRYTEQRAQEWLEARNRDGYLYVEVAPIYNADLIPRAV----- 193
DB 147 GAYGANDSNPEGMLYYENRLDSWLALHPDFWLDVKYTFIYSGNEVWFQIQLQYGVIDS 206
QY 194 -----VWSMQSDNTINEK-----VLVYNTANGYTYINHGTFQK 229
DB 207 SGEILLIRLNSNKESIDENGVTTVILENSAPNINDYLNGTATPK 251

```

Db 99 FNPPGW-----H-NYKLT-DANGKT-----TWLMDRGHLVGYQFSGLNDEPK-NLVTMTKYL 147  
 Qy 146 NVGRDQKG-GMYTEQRAQEWLEARNRDGYLYYEVAPIYNADELIPRAVVVSMQSSD 201  
 Db 148 NTGFSKDNPLGMLYYENRLDSWLALHPNFWDIKYTPVTHKNELVPRQVLYVYVIGD 204  
 RESULT 11  
 ABB53460  
 ID ABB53460 standard; Protein; 263 AA.  
 XX AC ABB53460;  
 XX AC 16-MAY-2002 (first entry)  
 XX DE Lactococcus lactis protein ybfb.  
 XX KW Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.  
 XX OS Lactococcus lactis IL1403.  
 XX PN FR2807446-A1.  
 XX PD 12-OCT-2001.  
 XX PF 11-APR-2000; 2000FR-0004630.  
 XX PR 11-APR-2000; 2000FR-0004630.  
 XX PA (INRG) INRA INST NAT RECH AGRONOMIQUE.  
 XX PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;  
 XX DR WPI; 2002-043418/06.  
 XX PT New nucleotide sequence useful in the identification of Lactococcus  
 PT lactis and related species -  
 PS Claim 6; SEQ ID No 162; 2504pp; French.  
 XX CC The present invention is related to a Lactococcus lactis nucleotide  
 CC sequence (ABA90521) and related proteins (ABB53300-ABB55621). The  
 CC nucleic acid sequence is useful in the detection and/or amplification of  
 CC nucleic acid sequence, particularly to identify Lactococcus lactis or  
 CC related species. The proteins of the invention are useful for the  
 CC biosynthesis or biodegradation of a composition of interest. The  
 CC invention helps research in lactic bacteria, particularly useful in the  
 CC production of yogurt and cheese.  
 CC Note: The sequence data for this patent is based on equivalent patent  
 CC WO200177334 (published 18-OCT-2001) which is available in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX SQ Sequence 263 AA;  
 Query Match 12.6%; Score 153.5; DB 23; Length 263;  
 Best Local Similarity 24.4%; Pred. No. 1.4e-07;  
 Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;  
 Qy 26 TFNDSPNYKTLGTS-QITPALFPKAGDILYSLKLDLGR-----RTARGTL 71  
 Db 39 TDNSSQVSTKSLASSVYKQAPLTFKNQRMVMTANTDALGRAVDVSHIQLKDSQEPKVKREPL 98  
 Qy 72 TYANVEGSGVGRQSGFKNQNPAGTGNPNHVKYKIEWLNGLSVYGVDFW---NRSHLIA--- 126  
 Db 99 TY-----NPGW-----H-NYFYKKSDDSGIGKMWLMARGHLVGYQF 135  
 Qy 127 DSLGGDALRY-----NAVGTGRTQNVGGDQKGRMYTEQRAQEWLEARNRDGYLYY 177  
 Db 136 SGLNNEARNLVPTAWFNCGNFTGTNDGNT-----ASMLYENRLDSWLALHPNYLYDY 189  
 Qy 178 EVAPIYNADELIPRAV---VSMQSSDNTINEK-----VLVYNTANGYTI 219  
 Db 190 QVTFPLYEGNELLPQIRLAYVGDKNQOTLSIKLGGREKSGNGGATVVVLDNVAPNAKI 249

Db 99 FNPPGW-----H-NYKLT-DANGKT-----TWLMDRGHLVGYQFSGLNDEPK-NLVTMTKYL 147  
 Qy 146 NVGRDQKG-GMYTEQRAQEWLEARNRDGYLYYEVAPIYNADELIPRAVVVSMQSSD 201  
 Db 148 NTGFSKDNPLGMLYYENRLDSWLALHPNFWDIKYTPVTHKNELVPRQVLYVYVIGD 204  
 RESULT 10  
 ABP25596  
 ID ABP25596 standard; Protein; 268 AA.  
 XX AC ABP25596;  
 XX AC 02-JUL-2002 (first entry)  
 XX DE Streptococcus polypeptide SEQ ID NO 368.  
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus pyogenes.  
 XX PN W0200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB04789.  
 XX PR 27-OCT-2000; 2000GB-0026333.  
 XX PR 24-NOV-2000; 2000GB-0028727.  
 XX PR 07-MAR-2001; 2001GB-0005640.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.  
 XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 XX PI Tettelin H;  
 XX DR WPI; 2002-352536/38.  
 XX DR N-PSDB; ABN66227.  
 XX PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 PS Claim 1; Page 3190; 4525pp; English.  
 XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC Streptococcus/GAS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 XX SQ Sequence 268 AA;  
 Query Match 13.0%; Score 158.5; DB 23; Length 268;  
 Best Local Similarity 31.1%; Pred. No. 4.3e-08;  
 Matches 55; Conservative 26; Mismatches 65; Indels 31; Gaps 12;  
 Qy 37 LGTSQI-TPALFPKAG--DILYSKLDLGRTRTARGTLTYANV-----EGSYGVGRQSGKN 89  
 Db 47 LGTKATETPGILPTGYSQVLGDLNLRP-----TFHQLKQDDEPNIKR-GLX 98  
 Qy 90 QNPAGTGNPNHVKYKIEWLNGLSVYGVDFW---NRSHLIA--DALRNAVTTGRTQ 145

QY 220 NYHNGT 225  
 Db 250 NYADGT 255

## RESULT 12

AA91338  
 ID AA91338 standard; Protein; 262 AA.

XX AC AA91338;  
 XX DT 30-MAY-2000 (first entry)  
 XX DE Group B Streptococcus protein sequence SEQ ID NO:71.  
 XX KW Group B Streptococcus; Streptococcus agalactiae; protein antigen;  
 KW vaccine; screening; immunogen; detection; diagnosis; infection;  
 KW antibody; affibody; antibacterial.  
 XX OS Streptococcus agalactiae.

XX WO200006736-A2.  
 XX PD 10-FEB-2000.

XX PF 27-JUL-1999; 99WO-GB02444.  
 XX PR 27-JUL-1998; 98GB-0016335.  
 XX PR 19-MAR-1999; 99US-0125163.

XX PA (MICR-) MICROBIAL TECHNIQS LTD.  
 XX PI Le Page RWF, Wells JM, Hanniffy SB;  
 XX DR WPI; 2000-195299/17.

XX PT New Group B Streptococcus protein, useful as vaccine, for diagnosis of  
 PT Streptococcal infections and for screening of antibodies or affibodies

XX PS Claim 1; Fig 1; 123pp; English.  
 XX CC AAA05803 to AAA05872 encode proteins, polypeptides and peptides (given  
 CC in AA91275 to AA91343) isolated from Group B Streptococcus (GBS), also  
 CC known as Streptococcus agalactiae. The GBS polynucleotides and  
 CC polypeptides have antibacterial activity. Immunogenic compositions  
 CC comprising GBS polynucleotides or polypeptides can be used as vaccines  
 CC and for the treatment or prophylaxis of GBS infection. The  
 CC polynucleotides and polypeptides can also be used in the detection of GBS  
 CC and for screening DNA encoding bacterial cell envelope associated or  
 CC secreted antigens in gram positive bacteria. AAA05873 to AAA05941  
 CC represent primers used in the exemplification of the present invention.

XX SQ Sequence 262 AA;  
 Query Match 11.0%; Score 134.5; DB 21; Length 262;  
 Best Local Similarity 25.8%; Pred. No. 1.4e-05;  
 Matches 39; Conservative 25; Mismatches 60; Indels 27; Gaps 5;

QY 89 NQNPAGH--TGNPNHVYKIEWNLGLSVGDFWNRSHLIADSLGGDLAR-----VN 137  
 Db 125 NWKPLGHQVATNDHYGHAVD-----KGLIAYALAGNPNFKGMDASVSNPQN 171  
 QY 138 AVTGTRTQNGGRDQKGMRYTEQRAQEWLEARNRDLGYLYVEVAPIYNAD-ELIPRAVVVS 196  
 Db 172 VVTQTAHSNOSQKINGQNYEYSLVRKAVDQNK--RVRYRVTPLYRNNDTDLVPFAHLE 229  
 QY 197 MQSSDNTINEKVLVNTANGTYNYHNGTPT 227  
 Db 230 AKSDGDTLEFNVAIPNTQASTYMDYATGEIT 260

## RESULT 13

ABP27392

XX ID ABP27392 standard; Protein; 354 AA.

XX AC ABP27392;  
 XX DT 02-JUL-2002 (first entry)

XX DE Streptococcus polypeptide SEQ ID NO 3960.

XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus agalactiae.

XX WO200234771-A2.  
 XX PD 02-MAY-2002.

XX PF 29-OCT-2001; 2001WO-GB04789.  
 XX PR 27-OCT-2000; 2000GB-0026333.  
 XX PR 24-NOV-2000; 2000GB-0028727.  
 XX PR 07-MAR-2001; 2001GB-0005640.

XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.

XX PI Telford J, Massignani V, Margarit Ros VI, Grandi G, Fraser C;  
 PI Tettelin H;

XX DR WPI; 2002-352536/38.  
 XX DR N-PSDB; ABN68023.

XX PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX PS Claim 1; Page 3550; 4525pp; English.  
 XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (II), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.

XX SQ Sequence 354 AA;  
 Query Match 11.0%; Score 134; DB 23; Length 354;  
 Best Local Similarity 25.2%; Pred. No. 2.5e-05;  
 Matches 56; Conservative 32; Mismatches 82; Indels 52; Gaps 12;

QY 34 YKTLGTSQITPALFPKAGDILYSKLDELGRTRTARGTLTYANVEGSGVGRSQSGKNQ--- 90  
 Db 41 YKTKENGTELP---FKKQRQLVLGLDLDKGRA-----TFAHIQ--LKVDEPKKKVKR 88  
 QY 91 ---NPAGWTGNPNHVYKIEWNLGLSVGDFW--NRSHLIADSLGG--DALRVNAVVTGTRT 144  
 Db 89 LKTTTPVGW-----H-NPKFYINDGTQKA---WLMRGRLLCHQFSGLNNEKRLVPTNW 139  
 QY 145 QNVGGRDQKG-----GMRYTEQRAQEWLEARNRDLGYLYVEVAPIYNADLIPRAVVVSMQS 199

Db 140 LNTGNYASTNSNPESMLFYKQLTWLSTHKVYLDYKVTPIYQNNELPKRIELKYVG 199  
 QY 200 SDNT-----INKE-----VLVYNTANGYTYNHNG 224  
 Db 200 IDKTGKLLPIFIGNKSTQDQFGISTVTLENTSPNATIDYLSG 241

RESULT 14  
 ABP26062  
 ID ABP26062 standard; Protein; 157 AA.  
 XX AC ABP26062;  
 XX DT 02-JUL-2002 (first entry)  
 XX DE Streptococcus polypeptide SEQ ID NO 1300.  
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus agalactiae.  
 XX PN WC200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB04789.  
 XX PR 27-OCT-2000; 2000GB-0026333.  
 XX PR 24-NOV-2000; 2000GB-0028727.  
 XX PR 07-MAR-2001; 2001GB-0005640.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.  
 XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX WPI; 2002-352536/38.  
 DR N-PSDB; ABN66693.  
 XX PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX PS Claim 1; Page 3286; 4525pp; English.  
 XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 XX Sequence 157 AA;  
 Query Match 10.7%; Score 131; DB 23; Length 157;  
 Best Local Similarity 26.0%; Pred. No. 1.5e-05;  
 Matches 39; Conservative 25; Mismatches 60; Indels 26; Gaps 5;

QY 89 NONPAG--TGNPNHVKKIEWLNGLSYVGDFWNRSHLIADSLGSD-----ALRVNA 138  
 Db 21 NWKPLGMHQVATNDHYGHAVD-----KGHLIAYALAGNFKGMDASVSNPQV 67  
 QY 139 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRDGILYVEVAPIYNAD-ELIPRAVVVSM 197  
 Db 68 VTQTAHSNQSOKINRGQNYVESLVKRAVDQNK--RVRYRVTPLYRNDTDLVPFAMHLEA 125

QY 198 QSSDNTINEKVLVNTANGYTYNYHNGTPT 227  
 Db 126 KSDGTLEFNFVAIPNTQASVTMDYATGEIT 155

RESULT 15  
 ABP30074  
 ID ABP30074 standard; Protein; 261 AA.  
 XX AC ABP30074;  
 XX DT 02-JUL-2002 (first entry)  
 XX DE Streptococcus polypeptide SEQ ID NO 9324.  
 XX KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.  
 XX OS Streptococcus agalactiae.  
 XX PN WC200234771-A2.  
 XX PD 02-MAY-2002.  
 XX PF 29-OCT-2001; 2001WO-GB04789.  
 XX PR 27-OCT-2000; 2000GB-0026333.  
 XX PR 24-NOV-2000; 2000GB-0028727.  
 XX PR 07-MAR-2001; 2001GB-0005640.  
 XX PA (CHIR-) CHIRON SPA.  
 XX PA (GENO-) INST GENOMIC RES.  
 XX PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;  
 PI Tettelin H;  
 XX WPI; 2002-352536/38.  
 DR N-PSDB; ABN70705.  
 XX PT New Streptococcus protein for the treatment or prevention of infection  
 PT or disease caused by Streptococcus bacteria, such as meningitis, and  
 PT for detecting a compound that binds to the protein -  
 XX PS Claim 1; Page 4057; 4525pp; English.  
 XX CC The invention relates to a protein (ABP25413-ABP30895) from group B  
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
 CC the specification. The proteins have antibacterial and antiinflammatory  
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and  
 CC antibodies that bind (I) are used in the manufacture of medicaments for  
 CC the treatment or prevention of infection or disease caused by  
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
 CC biological sample. (I) is used to determine whether a compound binds to  
 CC used as a vaccine or diagnostic composition. The disease caused by  
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
 CC acid encoding (I) may be used to recombinantly produce (I) and may be  
 CC used in gene therapy. Antibodies to (I) are used for affinity  
 CC chromatography, immunoassays, and distinguishing/identifying  
 CC Streptococcus proteins.  
 XX Sequence 261 AA;

Query Match 10.7%; Score 131; DB 23; Length 261;  
 Best Local Similarity 26.0%; Pred. No. 3.3e-05;  
 Matches 39; Conservative 25; Mismatches 60; Indels 26; Gaps 5;

Qy	89	NONPAGW--TCNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGCD-----ALRVNA	138
Db	125	NWKPLGWHQVATNDHYGHVD-----KGHLIAYALAGNFKGWDASVSNPQNV	171
Qy	139	VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVAPIYNAD-ELIPRAVVVSM	197
Db	172	VTQTAHSNQSNQKINRGQNYYESLVKAVDQNK--RVRYRVTPLYRNDTDLVPPFAMHLEA	229
Qy	198	QSSDNTINEKVLVYNTANGTYTINYHNGTPT	227
Db	230	KSQDGLLEFNVAIPNTQASVTMDYATGEIT	259

Search completed: January 5, 2004, 18:38:09  
 Job time : 60.0424 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:35:50 ; Search time 26.1473 Seconds  
(without alignments)  
370.561 Million cell updates/sec

Title: US-08-482-785-9

Perfect score: 1220

Sequence: 1 RQTQVNDVLDGASKYLN.....VYNTANGYTYNHNGPTQK 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220	100.0	229	4	US-08-393-889-9
2	1220	100.0	229	5	PCT-US94-09450-9
3	1220	100.0	271	4	US-08-393-889-8
4	1220	100.0	271	5	PCT-US94-09450-8
5	1215	99.6	272	4	US-08-393-889-15
6	1215	99.6	272	5	PCT-US94-09450-15
7	202	16.6	38	4	US-08-393-889-6
8	202	16.6	38	5	PCT-US94-09450-6
9	170	13.9	32	4	US-08-393-889-16
10	170	13.9	32	5	PCT-US94-09450-16
11	113	9.3	97	4	US-09-107-532A-5038
12	106	8.7	23	4	US-08-393-889-4
13	106	8.7	23	5	PCT-US94-09450-4
14	95.5	7.8	379	4	US-09-107-532A-6904
15	91.5	7.5	869	2	US-08-483-101-15
16	91.5	7.5	1118	3	US-09-379-523-3
17	88.5	7.3	1008	4	US-09-308-453-2
18	88	7.2	2137	4	US-09-134-001C-4463
19	87.5	7.2	858	4	US-09-252-991A-29756
20	87	7.1	420	3	US-09-239-303-9
21	85.5	7.0	1185	4	US-09-134-001C-5276
22	85.5	7.0	1475	3	US-09-007-999-2
23	85.5	7.0	1475	3	US-09-210-361-2
24	85.5	7.0	1475	4	US-09-740-274-2
25	84	6.9	340	4	US-09-120-051D-46
26	84	6.9	621	4	US-09-311-626B-2
27	83.5	6.8	1198	4	US-09-199-637A-405

28	83	6.8	340	2	US-08-355-844-1	Sequence 1, Appli
29	83	6.8	340	5	PCT-US95-16126-1	Sequence 1, Appli
30	82.5	6.8	528	2	US-08-793-229-35	Sequence 35, Appli
31	82.5	6.8	528	3	US-09-285-957-35	Sequence 35, Appli
32	82.5	6.8	1298	4	US-09-252-991A-30579	Sequence 30579, A
33	82.5	6.8	1416	4	US-09-071-035-404	Sequence 404, App
34	82.5	6.8	1448	4	US-09-071-035-402	Sequence 402, App
35	81.5	6.7	609	4	US-09-198-452A-579	Sequence 579, App
36	81.5	6.7	1073	4	US-09-206-942-47	Sequence 49, Appli
37	81.5	6.7	1079	4	US-09-206-942-47	Sequence 47, Appli
38	81	6.6	443	3	US-09-120-365-69	Sequence 69, Appli
39	81	6.6	443	3	US-09-515-039-69	Sequence 69, Appli
40	81	6.6	666	2	US-08-737-716-14	Sequence 14, Appli
41	81	6.6	811	1	US-08-480-604A-7	Sequence 7, Appli
42	81	6.6	811	2	US-08-405-496A-7	Sequence 7, Appli
43	81	6.6	811	3	US-08-915-136-7	Sequence 7, Appli
44	81	6.6	811	4	US-08-957-310-7	Sequence 7, Appli
45	81	6.6	811	4	US-10-011-366-7	Sequence 7, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-393-889-9  
; Sequence 9, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patsy P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; US-08-393-889-9

Query Match 100.0%; Score 1220; DB 4; Length 229;  
Best Local Similarity 100.0%; Pred. No. 2.3e-122;

Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Db	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Qy	61	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	120						
Db	61	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	120						
Qy	121	RSHLIADSLGGDALRVNAVVTGRTQNVGRDQKGMRYTEQRAQEWLEARNR	180						
Db	121	RSHLIADSLGGDALRVNAVVTGRTQNVGRDQKGMRYTEQRAQEWLEARNR	180						
Qy	181	PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYTINHNGTPTQK	229						
Db	181	PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYTINHNGTPTQK	229						
RESULT 2									
PCT-US94-09450-9									
; Sequence 9, Application PC/TUS9409450									
; GENERAL INFORMATION:									
; APPLICANT: Beckman Instruments, Inc.									
; APPLICANT: 2500 Harbor Boulevard									
; APPLICANT: Fullerton, California									
; TITLE OF INVENTION: Recombinant DNase B Derived from									
; NUMBER OF SEQUENCES: 16									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: Beckman Instruments, Inc.									
; STREET: 2500 Harbor Boulevard									
; CITY: Fullerton									
; STATE: California									
; COUNTRY: USA									
; ZIP: 92634									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Floppy disk									
; COMPUTER: IBM PC compatible									
; OPERATING SYSTEM: PC-DOS/MS-DOS									
; SOFTWARE: Patent In Release #1.0, Version #1.25									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: PCT/US94/09450									
; FILING DATE: 18-AUG-1994									
; CLASSIFICATION:									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER: US 08/082,845									
; FILING DATE: 23-JUN-1993									
; ATTORNEY/AGENT INFORMATION:									
; NAME: May, William H.									
; REGISTRATION NUMBER: 26,769									
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: (818) 796-4000									
; TELEFAX: (818) 795-6321									
; INFORMATION FOR SEQ ID NO: 9:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 229 amino acids									
; TYPE: amino acid									
; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
; HYPOTHETICAL: NO									
; ORIGINAL SOURCE:									
; ORGANISM: Streptococcus pyogenes									
PCT-US94-09450-9									
Query Match 100.0%; Score 1220; DB 5; Length 229;									
Best Local Similarity 100.0%; Pred. No. 2.3e-122;									
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Db	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						

Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Db	1	RQTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Qy	61	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	120						
Db	61	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	120						
Qy	121	RSHLIADSLGGDALRVNAVVTGRTQNVGRDQKGMRYTEQRAQEWLEARNR	180						
Db	121	RSHLIADSLGGDALRVNAVVTGRTQNVGRDQKGMRYTEQRAQEWLEARNR	180						
Qy	181	PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYTINHNGTPTQK	229						
Db	181	PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYTINHNGTPTQK	229						
RESULT 3									
US-08-393-889-8									
; Sequence 8, Application US/08393889									
; Patent No. 6420152									
; GENERAL INFORMATION:									
; APPLICANT: Adams, Craig W.									
; APPLICANT: Pang, Patty P.-Y.									
; APPLICANT: Belci, Marina									
; TITLE OF INVENTION: Recombinant DNase B Derived from									
; NUMBER OF SEQUENCES: 16									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: Sheldon & Mak									
; STREET: 225 South Lake Avenue, Ninth Floor									
; CITY: Pasadena									
; STATE: California									
; COUNTRY: USA									
; ZIP: 91001									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Floppy disk									
; COMPUTER: IBM PC compatible									
; OPERATING SYSTEM: PC-DOS/MS-DOS									
; SOFTWARE: Patent In Release #1.0, Version #1.25									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/08/393,889									
; FILING DATE:									
; CLASSIFICATION: 435									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER: US/08/082,845									
; FILING DATE:									
; ATTORNEY/AGENT INFORMATION:									
; NAME: Farber, Michael B.									
; REGISTRATION NUMBER: 32,612									
; REFERENCE/DOCKET NUMBER: 9521									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: (818) 796-4000									
; TELEFAX: (818) 795-6321									
; INFORMATION FOR SEQ ID NO: 8:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 271 amino acids									
; TYPE: amino acid									
; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
; US-08-393-889-8									
Query Match 100.0%; Score 1220; DB 4; Length 271;									
Best Local Similarity 100.0%; Pred. No. 3e-122;									
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	60						
Db	43	ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPKAGDILYSKLD	102						
Qy	61	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	120						
Db	103	LGRTTRTARGLTYANVEGSYGVRSFGKNQNPAGTGNPNHVYKIEWLNGLSYVGD	162						
Qy	121	RSHLIADSLGGDALRVNAVVTGRTQNVGRDQKGMRYTEQRAQEWLEARNR	180						

Db 163 RSHLIADSLGDLARVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 222  
Qy 181 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 229  
Db 223 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271

## RESULT 4

PCT-US94-09450-8  
; Sequence 8, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Beckman Instruments, Inc.  
; STREET: 2500 Harbor Boulevard  
; CITY: Fullerton  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92634  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/09450  
; FILING DATE: 18-AUG-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/082,845  
; FILING DATE: 23-JUN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: May, William H.  
; REGISTRATION NUMBER: 26,769  
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-09450-8

Query Match 100.0%; Score 1220; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPPKAGDILYSKLDE 60  
Db 43 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPPKAGDILYSKLDE 102  
Qy 61 LGRTRTARGTLTYANVSGSYGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFFWN 120  
Db 103 LGRTRTARGTLTYANVSGSYGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFFWN 162  
Qy 121 RSHLIADSLGDLARVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 180  
Db 163 RSHLIADSLGDLARVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 222  
Qy 181 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 229  
Db 223 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 271

## RESULT 5

US-08-393-889-15  
; Sequence 15, Application US/08393889  
; Patent No. 6420152  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belsi, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-393-889-15

Query Match 99.6%; Score 1215; DB 4; Length 272;  
Best Local Similarity 100.0%; Pred. No. 1e-121;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2 QTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPPKAGDILYSKLDEL 61  
Db 45 QTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALPPKAGDILYSKLDEL 104  
Qy 62 GRTRTARGTLTYANVSGSYGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFFWN 121  
Db 105 GRTRTARGTLTYANVSGSYGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFFWN 164  
Qy 122 SHLIADSLGDLARVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 181  
Db 165 SHLIADSLGDLARVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDLGYLYEVA 224  
Qy 182 IYNADLIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 229  
Db 225 IYNADLIPRAVVVSMQSSDNTINEKVLVYNTANGTYTINYHNGTPTQK 272

## RESULT 6

PCT-US94-09450-15  
; Sequence 15, Application PC/TUS9409450  
; GENERAL INFORMATION:  
; APPLICANT: Beckman Instruments, Inc.  
; APPLICANT: 2500 Harbor Boulevard  
; APPLICANT: Fullerton, California 92634  
; TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Beckman Instruments, Inc.

STREET: 2500 Harbor Boulevard

CITY: Fullerton

STATE: California

COUNTRY: USA

ZIP: 92634

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09450

FILING DATE: 18-AUG-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/082,845

FILING DATE: 23-JUN-1993

ATTORNEY/AGENT INFORMATION:

NAME: May, William H.

REGISTRATION NUMBER: 26,769

REFERENCE/DOCKET NUMBER: 39D-1357 PCT

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 272 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

PCT-US94-09450-15

Query Match

Best Local Similarity 99.6%; Score 1215; DB 5; Length 272;

Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QTVNSNDVNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDL 61

Db 45 QTVNSNDVNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDL 104

QY 62 GRTTRAGTLTYANVEGSGVGRQSGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDFWNR 121

Db 105 GRTTRAGTLTYANVEGSGVGRQSGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDFWNR 164

QY 122 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQWLEARNRGGYLYEVAP 181

Db 165 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQWLEARNRGGYLYEVAP 224

QY 182 IYNADLIPRAVVMQSSDNTINEKVLVYNTANGYINHGPTOK 229

Db 225 IYNADLIPRAVVMQSSDNTINEKVLVYNTANGYINHGPTOK 272

RESULT 7

US-08-393-889-6

Sequence 6, Application US/08393889

Patent No. 6420152

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/393,889

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/082,845

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.

REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 38 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Streptococcus pyogenes

US-08-393-889-6

Query Match

Best Local Similarity 16.6%; Score 202; DB 4; Length 38;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ROTQVNSNDVNDGASKYLNEALAWTFNDSPNYKTLG 38

Db 1 ROTQVNSNDVNDGASKYLNEALAWTFNDSPNYKTLG 38

RESULT 8

PCT-US94-09450-6

Sequence 6, Application PC/TUS9409450

GENERAL INFORMATION:

APPLICANT: Beckman Instruments, Inc.

APPLICANT: 2500 Harbor Boulevard

APPLICANT: Fullerton, California 92634

TITLE OF INVENTION: Recombinant DNase B Derived from

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Beckman Instruments, Inc.

STREET: 2500 Harbor Boulevard

CITY: Fullerton

STATE: California

COUNTRY: USA

ZIP: 92634

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/09450

FILING DATE: 18-AUG-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/082,845

FILING DATE: 23-JUN-1993

ATTORNEY/AGENT INFORMATION:

NAME: May, William H.

REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
PCT-US94-09450-6

Query Match 16.6%; Score 202; DB 5; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1e-14;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNYKTLG 38  
|||||  
Db 1 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNYKTLG 38  
|||||

RESULT 9  
US-08-393-889-16  
Sequence 16, Application US/08393889  
Patent No. 6420152  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-08-393-889-16

Query Match 13.9%; Score 170; DB 4; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.1e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNY 33  
|||||  
Db 1 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNY 32  
|||||

RESULT 10  
PCT-US94-09450-16  
Sequence 16, Application PC/TUS9409450  
GENERAL INFORMATION:  
APPLICANT: Beckman Instruments, Inc.  
APPLICANT: 2500 Harbor Boulevard  
APPLICANT: Fullerton, California 92634  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Beckman Instruments, Inc.  
STREET: 2500 Harbor Boulevard  
CITY: Fullerton  
STATE: California  
COUNTRY: USA  
ZIP: 92634  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09450  
FILING DATE: 18-AUG-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/082,845  
FILING DATE: 23-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: May, William H.  
REGISTRATION NUMBER: 26,769  
REFERENCE/DOCKET NUMBER: 39D-1357 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
PCT-US94-09450-16

Query Match 13.9%; Score 170; DB 5; Length 32;  
Best Local Similarity 100.0%; Pred. No. 2.1e-11;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNY 33  
|||||  
Db 1 QTQVSNDDVVLNDGASKYLNEALAWTFNDSPNY 32  
|||||

RESULT 11  
US-09-107-532A-5038  
Sequence 5038, Application US/09107532A  
Patent No. 6563275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

; REGISTRATION NUMBER: 26,769
; REFERENCE/DOCKET NUMBER: 39D-1357 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; PCT-US94-09450-4

Query Match      8.7%; Score 106; DB 5; Length 23;
Best Local Similarity 95.7%; Pred. No. 8.8e-05;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 QTQVSNDDVVLNDGASKYLNEALA 24
        |||||||
DB      1 QTQVSNDDVVLNDGASKYLNEALA 23

RESULT 14
US-09-107-532A-6904
; Sequence 6904, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6904:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 379 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; NAME/KEY: misc feature

```

Best Local Similarity 22.3%; Pred. No. 0.7;  
 Matches 41; Conservative 32; Mismatches 46; Indels 65; Gaps 13;

Qy	75	NVEGSY-----GVROSF-----GKNQNPACWTGNPNHVYK-----IEWLN 110
Db	432	NAKGYQSI SYTDGFSLSFYHNDKRVDCGRNYN-AGWSG--CYESYSASLSIPLLGWT 488
Qy	111	GLSYVGDFWN---RSHLIADSLGGDALRVNAVTCRTQNV-GGRDOKGGMRYTEQRAQEW 166
Db	489	TLGYSDTYSSESVYKSHLSE-----YGFYNQNIYKGTOR--WOLTSSTSLKW 534
Qy	167	LEARNDCGLYYEVAPIYNAD--ELIPR-----AVVWSQSSDNTINEKVLVYNTANGYTIN 220
Db	535	MD-----YNWPAIGIYNSEQRLTDKGGYISVITITRASRENSLN-----TGYSYN 580
Qy	221	YHNG 224
Db	581	YSRG 584

Search completed: January 5, 2004, 18:44:02  
 Job time : 27.1473 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:41:44 ; Search time 424.262 Seconds  
(without alignments)  
108.298 Million cell updates/sec

Title: US-08-482-785-9

Perfect score: 1220

Sequence: 1 RQTVSDVNLGASKYLN.....VYNTANGYTYNHNGTPTQX 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1220	100.0	229	9	US-08-119-900-9
2	1220	100.0	271	9	US-08-119-900-8
3	1215	99.6	272	9	US-08-119-900-15
4	202	16.6	38	9	US-08-119-900-6
5	170	13.9	32	9	US-08-119-900-16
6	131	10.7	261	12	US-08-769-736-129
7	122.5	10.0	274	12	US-08-769-744A-168
8	106	8.7	23	9	US-08-119-900-4
9	106	8.7	1233	10	US-08-738-626-4312
10	97	8.0	2386	15	US-10-156-761-7751
11	92	7.5	867	10	US-08-839-894-6
12	90	7.4	398	12	US-10-369-493-3703
13	87.5	7.2	608	15	US-10-156-761-8616
14	87	7.1	422	12	US-10-369-493-1312
15	87	7.1	422	12	US-10-369-493-20377

16	87	7.1	428	12	US-10-284-400-8
17	87	7.1	429	12	US-10-284-400-16
18	86	7.0	226	12	US-10-029-386-34033
19	86	7.0	1136	9	US-08-815-242-12447
20	86	7.0	1136	9	US-08-815-242-12814
21	86	7.0	1179	9	US-08-815-242-5522
22	85.5	7.0	275	12	US-10-323-069A-105
23	85.5	7.0	1475	9	US-08-740-274-2
24	85	7.0	420	12	US-10-369-493-21586
25	85	7.0	522	12	US-10-172-502-17
26	85	7.0	1113	12	US-10-369-493-2198
27	84.5	6.9	540	12	US-10-369-493-3703
28	84	6.9	362	12	US-10-369-493-786
29	84	6.9	621	15	US-10-124-880-2
30	83.5	6.8	1198	11	US-08-975-719-405
31	83	6.8	227	12	US-10-237-386-27
32	83	6.8	398	9	US-08-815-242-10390
33	83	6.8	398	12	US-10-369-493-23607
34	83	6.8	538	12	US-10-369-493-11935
35	83	6.8	1156	9	US-08-815-242-13187
36	83	6.8	1436	15	US-10-080-505-13
37	82.5	6.8	1795	7	US-08-973-363-17
38	82	6.7	398	12	US-10-369-493-21321
39	81.5	6.7	1073	12	US-10-193-764-45
40	81.5	6.7	1079	12	US-10-193-764-43
41	81.5	6.7	1723	9	US-08-841-132-394
42	81.5	6.7	1723	9	US-08-841-132-395
43	81	6.6	479	10	US-08-881-752A-14
44	81	6.6	534	12	US-10-369-493-16599
45	81	6.6	811	12	US-10-354-774-7

#### ALIGNMENTS

#### RESULT 1

US-09-119-900-9  
; Sequence 9, Application US/09119900  
; Patent No. US20020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:

Sequence 8, Appli  
Sequence 16, Appli  
Sequence 34033, A  
Sequence 12447, A  
Sequence 12814, A  
Sequence 5522, Ap  
Sequence 105, App  
Sequence 2, Appli  
Sequence 21586, A  
Sequence 17, Appl  
Sequence 2198, Ap  
Sequence 3703, Ap  
Sequence 786, App  
Sequence 405, App  
Sequence 2, Appli  
Sequence 27, Appl  
Sequence 10390, A  
Sequence 23607, A  
Sequence 11935, A  
Sequence 13187, A  
Sequence 17, Appl  
Sequence 21321, A  
Sequence 45, Appl  
Sequence 43, Appl  
Sequence 394, App  
Sequence 395, App  
Sequence 14, Appl  
Sequence 16599, A  
Sequence 7, Appli

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
US-09-119-900-9

Query Match 100.0%; Score 1220; DB 9; Length 229;
Best Local Similarity 100.0%; Pred. No. 1.1e-117;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQVSNVNDGASKYLNEALAWTFNDSNYYKTLGTSQITPALPPKAGDILYSKLE 60
Db 1 RQTQVSNVNDGASKYLNEALAWTFNDSNYYKTLGTSQITPALPPKAGDILYSKLE 60
Qy 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSYVGDFWN 120
Db 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSYVGDFWN 120
Qy 121 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180
Db 121 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180
Qy 181 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYTINYHNGTPTQK 229
Db 181 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYTINYHNGTPTQK 229

RESULT 2
US-09-119-900-8
; Sequence 8, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-119-900-15

Query Match 99.6%; Score 1215; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.5e-117;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-119-900-8

Query Match 100.0%; Score 1220; DB 9; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.4e-117;
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQVSNVNDGASKYLNEALAWTFNDSNYYKTLGTSQITPALPPKAGDILYSKLE 60
Db 43 RQTQVSNVNDGASKYLNEALAWTFNDSNYYKTLGTSQITPALPPKAGDILYSKLE 102
Qy 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSYVGDFWN 120
Db 103 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSYVGDFWN 162
Qy 121 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180
Db 163 RSHLIADSLGGDALRVNAVTTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 222
Qy 181 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYTINYHNGTPTQK 229
Db 223 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYTINYHNGTPTQK 271

RESULT 3
US-09-119-900-15
; Sequence 15, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-119-900-15

Query Match 99.6%; Score 1215; DB 9; Length 272;
Best Local Similarity 100.0%; Pred. No. 4.5e-117;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 QTOVSNVNDGASKYLNEALAWTFNDSNNYKTLGTSGITPALKPKAGDILYSLKDEL 61  
DB 45 QTOVSNVNDGASKYLNEALAWTFNDSNNYKTLGTSGITPALKPKAGDILYSLKDEL 104  
QY 62 GTRTRTARGTLTYANVEGSGVROSGFNQNPAGWTGNPNHVKYKIEWLNGLSYVGDFFWR 121  
DB 105 GTRTRTARGTLTYANVEGSGVROSGFNQNPAGWTGNPNHVKYKIEWLNGLSYVGDFFWR 164  
QY 122 SHLIADSLGDLARVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDSGLYLYEYVAP 181  
DB 165 SHLIADSLGDLARVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDSGLYLYEYVAP 224  
QY 182 IYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGYTYNHNGTPTQK 229  
DB 225 IYNADELIPRAVVVSMQSSDNTINEKVLVYNTANGYTYNHNGTPTQK 272

## RESULT 4

US-09-119-900-6  
; Sequence 6, Application US/09119900  
; Patent No. US20020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; US-09-119-900-6

Query Match 16.6%; Score 202; DB 9; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ROTQVSNVNDGASKYLNEALAWTFNDSNNYKTLG 38  
DB 1 ROTQVSNVNDGASKYLNEALAWTFNDSNNYKTLG 38

## RESULT 5

US-09-119-900-16  
; Sequence 16, Application US/09119900  
; Patent No. US20020081622A1  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; US-09-119-900-16

Query Match 13.9%; Score 170; DB 9; Length 32;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QTOVSNVNDGASKYLNEALAWTFNDSNNY 33  
DB 1 QTOVSNVNDGASKYLNEALAWTFNDSNNY 32

## RESULT 6

US-09-769-736-129  
; Sequence 129, Application US/09769736  
; Publication No. US20030138775A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Le Page, Richard WF  
; APPLICANT: Wells, Jeremy M  
; APPLICANT: Hanniffy, Sean B  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21089wo  
; CURRENT APPLICATION NUMBER: US/09/769,736  
; CURRENT FILING DATE: 2003-02-14  
; PRIOR APPLICATION NUMBER: GB 9816335.5

```
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125163
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 212
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 129
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
US-09-769-736-129

Query Match      10.7%; Score 131; DB 12; Length 261;
Best Local Similarity 26.0%; Pred. No. 3.5e-05;
Matches 39; Conservative 25; Mismatches 60; Indels 26; Gaps 5;

Qy 89 NQNPAGW--TGNPHVKYKIEWLNGLSVGVDFWNRSHLIADSLGDD-----ALRVNA 138
Db 125 NWKPLGHQVATNDHYGHAVD-----KGHLIAYALAGNFKGMDASVSNPQNV 171
Qy 139 VTGTRTQNVGGRDQKGGWRYTEQRAQEWLEARNRGGYLYEYEVAPIYNAD-ELIPRAVVVSM 197
Db 172 VTQTAHNSQSNQKINRGQNYTESLVKAVDQNK--RVRYRVTPLYRNDTDLVPPFAMHLEA 229
Qy 198 QSSDNTINEKVLVYNTANGYTYNTYHNGTPT 227
Db 230 KSDQTLEFNVAIENTQASYTMDYATGEIT 259

RESULT 7
US-09-769-744A-168
; Sequence 168, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 168
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-168

Query Match      10.0%; Score 122.5; DB 12; Length 274;
Best Local Similarity 22.3%; Pred. No. 0.00028;
Matches 57; Conservative 34; Mismatches 100; Indels 65; Gaps 11;

Qy 3  TQVSNVNDVLDGASKYLNEALWTFND-----SPNYKTLGTSTQI 42
Db 50  SQALAESVLTDAVKSQIKGSLEWNGSGAFVNGNKTNLDAKVSFKPVADNKTIVGKETV 109
Qy 43  TPALFPKAGDILYSKLEB--LGRTRTARGTLTYANVSGSVGVROSGKQNPAGWGTGNPN 100
Db 110 -----PTVANALLSKATQYKXKTKGTGTSW-----TPPGW----- 142
Qy 101  HVKYKIEWLNGLSVGVDFWNRSHLIADSL-----GGDALR---VNAVGTGRTQNVGGRDQ 152
Db 143  ---HGVKNLKG-SYTHAV-DRGHLGLYALIGLDGFDASTNPKNIAVQTAWANQAQAEY 197
Qy 153  KGMWRYTEQRAQEWLEARNRGGYLYEYEVAPIYNAD-ELIPRAVVVSMOSSDNTINEKVLVY 211
```

```
Db 198  STGQNYYESKVRKALDQNK--RVRYRVTLTYASNEDLVPSASQIEAKSSDGELEFNVLPV 255
Qy 212  NTANGYTINYHNGTPT 227
Db 256  NVQKGLQLDYRTGEVT 271

RESULT 8
US-09-119-900-4
; Sequence 4, Application US/09119900
; Patent No. US20020081622A1
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
US-09-119-900-4

Query Match      8.7%; Score 106; DB 9; Length 23;
Best Local Similarity 95.7%; Pred. No. 0.00043;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2  QTOVSNVNDVLDGASKYLNEALA 24
Db 1  QTOVSNVNDVLDGASKYLNEALA 23

RESULT 9
US-09-738-626-4312
; Sequence 4312, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
```

```
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 4312
; LENGTH: 1233
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4312

Query Match      8.7%; Score 106; DB 10; Length 1233;
Best Local Similarity 22.8%; Pred. No. 0.12;
Matches 55; Conservative 33; Mismatches 93; Indels 60; Gaps 10;

Qy 18 YLNEALAWTFNDSPNYKTLGTSQITPALPKAGDILYKSLDELGRTRTARGTLYANV- 76
Db 1020 YETBEAIERNSOSYIREIGGSEL-----WNILKGNSEGLSLAQKCAPQATEINVI 1071
Qy 77 --EGSYGVRSFGKQNPAGWTGNPNHVKYKIEWLNG-----LSYVGDFWNRSHLI 125
Db 1072 RNSGLEAMRNLGADQSAEISADSRLRAQLEWNRRENLDLGQLPTLFDPAEKYEYL 1131
Qy 126 ADSLGGBALRVNAVGTGRTQNVGRDQKGNRYTEQRAQEWLEARNRQGYLY-----Y 177
Db 1132 IDHLGDDRIK---VTARELSLASEHRRG-----NAENWLAPYVSYSLNRMIAH 1181
Qy 178 EV-API-----VNADELIPRAVVSMQSSDNTINEKVLVYNTANGYTINYHNGTP 226
Db 1182 EVIRPIAQINYSRHDWANAARLIPRLTGFOLVSAE-----AKVL-----SAINNNNIIP 1230
Qy 227 T 227
Db 1231 T 1231

RESULT 10
US-10-156-761-7751
; Sequence 7751, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7751
; LENGTH: 2386
```

```
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-7751

Query Match      8.0%; Score 97; DB 15; Length 2386;
Best Local Similarity 21.7%; Pred. No. 2.6;
Matches 60; Conservative 25; Mismatches 108; Indels 84; Gaps 10;

Qy 3 TOVSNV--VLNDGASKY-----LNBALAWTFNDSPNYKTLGTSQITPALPKAG 51
Db 1433 TATTNDVQTVLTGYDNLSRATSVRSAGADELAAVWDDPA---ATGGKGQITSVSRDAS 1489
Qy 52 DILYS-----KLDELGRTRTARGTL--TYANVESYGVRSFGKQNPAGWTGNPNHVKYK 105
Db 1490 GNTYTTKTGKFERGRPLNTTVTLPTVNGLAGDYTTSTYD-----A 1532
Qy 106 IEWLNGLSYVG-----DFWNRSHLADSLGGDALRVNAV----- 139
Db 1533 ADHITSVSYPAAGKLAAEKVTTTYDDYQGFRTLTSSLGGTAYIDNTTYDAYGLRVERDYG 1592
Qy 140 ----TGTRTQNVGRDQKGNRYTEQRAQEWL-----BANRDGYLYYEVAPIYNAD 188
Db 1593 AEFGGNGIQARQVGYDDSNGTNRWLSIAATTTTINDLVSEAQKDTYLYDNTGKLTRE- 1651
Qy 189 IPRAVVSMQSSDNTINEKVLVYNTANGYTINYHNGT 225
Db 1652 -----QASGQTAQSQCLRYDDQSRLTAYTHTT 1679

RESULT 11
US-09-839-894-6
; Sequence 6, Application US/09839894
; Patent No. US20020176868A1
; GENERAL INFORMATION:
; APPLICANT: Alboum, Zeev
; APPLICANT: Barry, Eileen M.
; APPLICANT: Levine, Myron M.
; APPLICANT: University of Maryland
; TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF THE
; TITLE OF INVENTION: CSA OPERON
; FILE REFERENCE: UOPMD.006A
; CURRENT APPLICATION NUMBER: US/09/839,894
; CURRENT FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/198,626
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 867
; TYPE: PRT
; ORGANISM: E. coli
; US-09-839-894-6

Query Match      7.5%; Score 92; DB 10; Length 867;
Best Local Similarity 20.4%; Pred. No. 2;
Matches 52; Conservative 43; Mismatches 68; Indels 92; Gaps 16;

Qy 5 VSNVDVLNDGASKYLNALAWTFNDSPNYKTLGTSQITPALPKAGDILYKSLDELGR 64
Db 385 ITKNISVQOGASVI-----DNKNYE-----GSLKWNSSGILSGSL 419
Qy 65 RTARGTLTYANVEGSY-----GVRQSF-----GKQNPAGWTGNPNHVKYK-- 105
Db 420 NSEFSLMGDNAGKNYQSIYSYTDGFSLSFVHNDKRVNDCGRNYN-AGWSG--CYESYSAS 476
Qy 106 -----IEWLNGLSYVGDFWNR-----HLIADSLGGDALRVNAVGTGRTQNV-GRDQKGG 155
Db 477 LSIPLLGWTSTLGY-SDTYSESVKKNHILSE-----YGFYQNIYKGRTOR-- 521
Qy 156 MRYTEQRAQEWLEARNRQGYLYYEVAPIYNAD--ELIPR-----AVVSMQSSDNTINEKVL 209
Db 522 WQLTSTSLKWD-----YNFMPAIGYNSRQQLTKGQYISVTLTRASRENSLNA--- 573
```

0  
4  
0  
0  
4  
0  
-  
0  
0  
1  
0  
1  
0  
0

; Sequence 20377, Application US/

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 17:31:14 ; Search time 26.5691 Seconds  
(without alignments)  
828.882 Million cell updates/sec

Title: US-08-482-785-9  
Perfect score: 1220  
Sequence: 1 RQTVSNDDVNLGASKYL.....VYNTANGYTYNHNGTPTQK 229

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1216	99.7	271	2	mitogenic factor,
2	365	29.9	327	2	deoxyribonuclease
3	153.5	12.6	263	2	hypothetical prote
4	122.5	10.0	274	2	DNA-entry nuclease
5	122.5	10.0	274	2	endA protein - Str
6	122.5	10.0	274	2	deoxyribonuclease
7	104.5	8.6	455	2	probable outer mem
8	94	7.7	319	2	alpha-toxin precu
9	91.5	7.5	869	2	cfaC protein precu
10	91.5	7.5	1534	2	P-glycoprotein - T
11	91	7.5	386	2	conserved hypotet
12	91	7.5	1203	2	pullulanase - Ther
13	90.5	7.4	292	2	botulinum neurotox
14	90	7.4	319	2	Alpha-Hemolysin pr
15	89.5	7.3	347	2	uncharacterized con
16	89	7.3	928	2	hypothetical prote
17	89	7.3	4199	2	hypothetical prote
18	88	7.2	774	2	autolysin (amidase
19	88	7.2	802	2	penicillin amidase
20	87	7.1	420	1	glutamate dehydrog
21	87	7.1	422	2	probable glutamate
22	87	7.1	809	2	hyaluronate lyase
23	86.5	7.1	611	2	hypothetical prote
24	86.5	7.1	1116	2	DNA-directed RNA p
25	86	7.0	396	2	alpha-galactosidas
26	86	7.0	729	2	hypothetical prote
27	86	7.0	1182	2	DNA-directed RNA p
28	86	7.0	1183	2	RNA polymerase bet
29	85.5	7.0	1475	2	gtfB protein precu

30 85.5 7.0 2044 2 AB1180 probable peptidogl  
31 85 7.0 362 2 D90755 outer membrane pro  
32 85 7.0 362 2 B85619 outer membrane pro  
33 85 7.0 420 2 D75176 glutamate dehydrog  
34 85 7.0 479 2 H71942 probable outer mem  
35 85 7.0 503 2 S31940 starch-degrading e  
36 85 7.0 725 2 C84423 probable ABC trans  
37 85 7.0 1165 2 T38867 probable DNA-direc  
38 85 7.0 1640 2 D86798 prophage p13 prote  
39 84 6.9 293 2 H44644 botulinum neurotox  
40 84 6.9 362 1 MMECF outer membrane por  
41 84 6.9 417 2 G70033 maltose/maltodextr  
42 84 6.9 836 2 A98014 hypothetical prote  
43 84 6.9 836 2 C85858 hypothetical prote  
44 84 6.9 1102 2 S77518 DNA-directed RNA p  
45 83.5 6.8 386 2 AG1120 conserved hypotet

## ALIGNMENTS

## RESULT 1

S36907 mitogenic factor, 25K, precursor - Streptococcus pyogenes

C:Species: Streptococcus Pyogenes

C:Date: 10-Dec-1993 #sequence revision 23-Feb-1996 #text\_change 15-Oct-1999

C:Accession: S36907; S36908; S29188

R:Iwasaki, M.; Igarashi, H.; Hinuma, Y.; Yutsudo, T.

FEBS Lett. 331, 187-192, 1993

A:Title: Cloning, characterization and overexpression of a Streptococcus pyogenes gene

A:Reference number: S36907; MUID:94009636; PMID:8405402

A:Accession: S36907

A:Molecule type: DNA

A:Residues: 1-271 <IWA>

A:Cross-references: EMBL:D13428; NID:g432369; PIDN:BA02693.1; PID:d1003198; PID:g432370

A:Accession: S36908

A:Molecule type: protein

A:Residues: 44-64 <IWA>

R:Yutsudo, T.; Murai, H.; Gonzalez, J.; Takao, T.; Shimonishi, Y.; Takeda, Y.; Igarashi,

FEBS Lett. 308, 30-34, 1992

A:Title: A new type of mitogenic factor produced by Streptococcus pyogenes.

A:Reference number: S29188; MUID:92354761; PMID:1644200

A:Accession: S29188

A:Molecule type: protein

A:Residues: 44-64 <YUT>

F:1-43/Domain: signal sequence #status predicted <SIG>

F:44-271/Product: mitogenic factor, 25K #status experimental <MAT>

Query Match 99.7%; Score 1216; DB 2; Length 271;  
Best Local Similarity 99.6%; Pred. No. 4.3e-96;  
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RQTVSNDDVNLGASKYLNEALAWTENDSPNYKTLGTSGITPALPKAGDILYSLKDE 60

DB 43 RQTVSNDDVNLGASKYLNEALAWTENDSPNYKTLGTSGITPALPKAGDILYSLKDE 102

QY 61 LGRTTRTARGTLTYANVEGSGVQSGFQKQNPAGTGNPNHVKYKIEWLNGLSVGDFWN 120

DB 103 LGRTTRTARGTLTYANVEGSGVQSGFQKQNPAGTGNPNHVKYKIEWLNGLSVGDFWN 162

QY 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 180

DB 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 222

QY 181 PIYNADLIPRAVWVSMQSSDNTINEKVLVYNTANGYTYNHNGTPTQK 229

DB 223 PIYNADLIPRAVWVSMQSSDNTINEKVLVYNTANGYTYNHNGTPTQK 271

## RESULT 2

JT0584 deoxyribonuclease I (EC 3.1.21.1) sdc precursor - Streptococcus "equisimilis"

N;Alternate names: streptodornase

C:Species: Streptococcus "equisimilis"  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 15-Oct-1999  
C:Accession: J0584; S24204  
R:Wolinowka, R.; Ceglowski, P.; Kok, J.; Venema, G.  
Gene 106, 115-119, 1991  
A:Title: Isolation, sequence and expression in Escherichia coli, Bacillus subtilis and L.  
A:Reference number: J0584; MUID:92039051; PMID:1937032  
A:Accession: J0584  
A:Molecule type: DNA  
A:Residues: 1-327 <WOL>  
A:Cross-references: ENBL.X17241; NID:G48693; PIDN:CAA35106.1; PID:G48694  
A:Experimental source: strain H46A  
C:Genetics:  
A:Gene: sdc  
C:Keywords: hydrolase  
F:1-35/Domain: signal sequence #status predicted <SIG>  
F:36-308/Product: deoxyribonuclease sdc #status predicted <MAT>  
Query Match 29.9%; Score 365; DB 2; Length 327;  
Best Local Similarity 34.6%; Pred. No. 1.6e-23;  
Matches 84; Conservative 34; Mismatches 79; Indels 46; Gaps 4;  
QY 26 TFNDSPNYKTLGTSQITPALFPK-----AGDILYSKLDLGRTR 65  
DB 43 TYGEYKDYTVIGESNIDQSAFPKLYKTERVYKQGTSEKRVTVSDVYVNPDLGKYST 102  
QY 66 TARGTLTYANVEGSGVRSQFQKQNPAGW-----TCN-----PNH 101  
DB 103 GAYGVTVKDMIDMSKGYREKWTNPEPSGWFYFNADNEISEKEYDSRRTKSKYVTNN 162  
QY 102 VKYKIEWNLGLSYVGDFFNRSHLIADSLGSDALRVNAVGTGRTONVGGDKGGRVYEQ 161  
DB 163 VPVVLTTLLKGYKNSHLFVASHLFDADSLGKSKIRKNAITGQMNVGTR--KGGHGYIEK 220  
QY 162 RAQEWEANRDGYLYVEVAPIYNADLIPRAVVSQSSDNTINEKVLVYNTANGYTINY 221  
DB 221 KVLSHITKPNPDVYVYSAIPEYQGAELLARSLVLSALSSDGVINETVAVENTADGFNIN 280  
QY 222 HNG 224  
DB 281 EKG 283  
RESULT 3  
D86644  
Hypothetical protein ybfB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
C:Species: Lactococcus lactis subsp. lactis  
C:Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
C:Accession: D86644  
R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich  
Genome Res. 11, 731-753, 2001  
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis se  
A:Reference number: A86625; MUID:21235186; PMID:11337471  
A:Accession: D86644  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-263 <STO>  
A:Cross-references: GB:AE005176; PID:gl2723006; PIDN:AAK04254.1; GSPDB:GN00146  
A:Experimental source: strain IL1403  
C:Genetics:  
A:Gene: ybfB  
Query Match 12.6%; Score 153.5; DB 2; Length 263;  
Best Local Similarity 24.4%; Pred. No. 1.2e-05;  
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;  
QY 26 TFNDSPNYKTLGTS-QITPALFPKAGDILYSKLDLGR-----RTARGTL 71  
DB 39 TDNSSQVTSKSLASSVKQAPLTFNQRMWANTDALGRAVDSHIQDSQBPVKREPL 98  
QY 72 TYANVEGSGVRSQFQKQNPAGTGNPNHVKYKIEWNLGLSYVGDFFW--NRSHLIA--- 126  
DB 99 TY-----NPVGW-----H-NYNFYKKS DSGSIGRQWLMARGHLVGYQF 135

QY 127 DSLGGDALRV-----NAVGTGRTONVGGDKGGRVYEQAEWLEARNDCYLYY 177  
DB 136 SGLNNEARNLVPETAFNGFNGTGTNDGT-----ASMLYYENLDSWLANHPNYLDY 189  
QY 178 EVAPIYNADLIPRAV--VVSQSSDNTINEK-----VLVYNTANGYTI 219  
DB 190 QVTPLYEGNELLPRQILRAYVIGDKNGOTLSIKLGGREKSGNGGATVVVLDNVAPNAKI 249  
QY 220 NYHNGT 225  
DB 250 NYADGT 255  
RESULT 4  
F95229  
DNA-entry nuclease [imported] - Streptococcus pneumoniae (strain TIGR)  
C:Species: Streptococcus pneumoniae  
C:Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:Accession: F95229  
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  
son, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001  
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916  
A:Accession: F95229  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <KUR>  
A:Cross-references: GB:AE005672; PIDN:AAK76031.1; PID:gl4973470; GSPDB:GN00164; TIGR:SP4  
A:Experimental source: strain TIGR4  
C:Genetics:  
A:Gene: SPI964  
Query Match 10.0%; Score 122.5; DB 2; Length 274;  
Best Local Similarity 22.3%; Pred. No. 0.0056;  
Matches 57; Conservative 34; Mismatches 100; Indels 65; Gaps 11;  
QY 3 TVNSNDVVLNDGASKYLNEALAWTFND-----SPNYYKTLGTSQI 42  
DB 50 SQALAESVLTDVAVKSIKGSLEWNGSGAFIVNGKTNLDKAVSKPYADNKTIVGKETV 109  
QY 43 TPALFPKAGDILYSKLDL--LGRTRTARGTLTYANVEGSGVRSQFQKQNPAGTGNPN 100  
DB 110 -----PTVANALLSKATROYKNEKGTGSGTSM-----TPPGW----- 142  
QY 101 HVKYKIEWNLGLSYVGDFFNRSHLIADSL-----GGDALR--VNAVGTGRTONVGGRDQ 152  
DB 143 ---HQVKNLKG-SYTHAV-DRGHLLGLYALIGLGDGFDASTNPKNTAVQTAWANQAQAY 197  
QY 153 KGGNRYTEQAEWLEARNRDGYLYVEVAPIYNADLIPRAVVSQSSDNTINEKVLVY 211  
DB 198 STGQNYYESKVRKALDQNK--RVRYRVTLTYASNEDLVPSASQIEAKSSDGELEFNVLP 255  
QY 212 NTANGYTYINHGNTPT 227  
DB 256 NVQKGLQLDYRTGVT 271  
RESULT 5  
SI0641  
endA protein - Streptococcus pneumoniae  
C:Species: Streptococcus pneumoniae  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 15-Oct-1999  
C:Accession: SI0641  
R:Puyet, A.; Greenberg, B.; Lacks, S.A.  
J. Mol. Biol. 213, 727-738, 1990  
A:Title: Genetic and structural characterization of endA. A membrane-bound nuclease requ  
A:Reference number: SI0640; MUID:90294291; PMID:2359120  
A:Accession: SI0641  
A:Status: preliminary

```

RESULT 7
G71896
probable outer membrane protein - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-May-2000
C;Accession: G71896
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric patho
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: G71896
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-455 <ARN>
A;Cross-references: GB:AE001503; GB:AE001439; NID:g4155275; PIDN:AAD06302.1; PID:g4155282
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0719
C;Superfamily: Helicobacter pylori hypothetical protein HP0209

Query Match      8.6%; Score 104.5; DB 2; Length 455;
Best Local Similarity 25.2%; Pred. NO. 0.37;
Matches 53; Conservative 19; Mismatches 85; Indels 53; Gaps 8;

QY      29 DSPNYKTLGTSQITP--ALFPKAGDIILY-----SKLDELGRTRTARGTLTYANVEGSY 80
DB      261 DSNPKFKGLGURAGTTINVIIPVYAKOLYDVYRNRSKIGEWGASLLIHORFDYNENFGF 320

QY      81 GVRQSFQGNQNPAGWTGNPNHVKYKIEWLNGLSYVGVDFWNRSHLIADSLGGDALRVNAV 140
DB      321 GYQNFQGNANARICWGNPIFPNVR---NNSVYGVGSN-----AITADAVS 364

QY      141 GTRTQNVGGRDQKGM-----RYT-EQRAQEWLENRDGYLYEVAPIYNADELIPRAVV 194
DB      365 G--YVFGGGVYRGFLGILGRYTVATRASERSINLNLGYKMGVSFARV----- 409

QY      195 VSMQSSDNTINEKVLVYNTANGYTYNHNG 224
DB      410 -----DVNLEYVYVSMHNGYRLDYLTG 431

RESULT 8
S69209
alpha-toxin precursor - Staphylococcus aureus
C;Species: Staphylococcus aureus
C;Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
C;Accession: S69209
R;Hedengrahn, G.
submitted to the EMBL Data Library, October 1992
A;Reference number: S69209
A;Accession: S69209
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-319 <HED>
A;Cross-references: EMBL:X01645; NID:g46763; PIDN:CAA25801.1; PID:g46765
C;Superfamily: leukocidin
C;Keywords: toxin
F;1-27/Domain: signal sequence #status predicted <SIG>
F;27-319/Product: alpha-toxin #status predicted <MAT>

Query Match      7.7%; Score 94; DB 2; Length 319;
Best Local Similarity 19.3%; Pred. NO. 1.8;
Matches 47; Conservative 42; Mismatches 88; Indels 66; Gaps 10;

QY      29 DSPNYKTLGTSQITPALFPKAGD-----ILYSKDELGRTRTARGTLTYAN 75
DB      28 DSDINIKT-GTTDIGSNVTYKTGDLVYDKENGMMHKVYFYSFIDDKNHNKLLVITKGT 86

QY      76 VEGSYGVRQSGFQGNQNPAGWTGNPNHVKYKIEWL--NGLSYVGDFWNRSHL-----IA 126

```

Db 87 IAGQYRVSEGANCKSLAW---PSAPKVLQLPDNEVAQISDYPPNSIDTKEYMSTLT 143  
Qy 127 DSLGGDALRVNAVVTGTRTQNVGG-----RDQKGMR--Y 158  
Db 144 YGFNGN-----VTGDDTKIGGLIGANVSIGHTLKYVPDFKFTILSPDTPDKVGMKVIF 197  
Qy 159 TEQRAQEWLEARNRGYLYVEVAPYINADDELIPRAVVVSMQSSDNTINEKVLVYNTANGYT 218  
Db 198 NMVNQNMGPYDRDSW-----NPVY-GNQLPMKTRNGSMKAADNFDLPNFKASSLLSSGFS 251  
Qy 219 INY 221  
Db 252 PDF 254

RESULT 9  
CS6617  
cfaC protein precursor - Escherichia coli plasmid NTP113  
C:Species: Escherichia coli  
C:Date: 05-Jan-1996 #sequence\_revision 05-Jan-1996 #text\_change 08-Oct-1999  
C:Accession: CS6617  
R:Jordi, B.J.; Willshaw, G.A.; van der Zeijst, B.A.; Gaastra, W.  
DNA Seq. 2, 257-263, 1992  
A:Title: The complete nucleotide sequence of region 1 of the CFA/I fimbrial operon of hu  
A:Reference number: A56617; MUID:92329981; PMID:1352712  
A:Accession: CS6617  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-869 <JOR>  
A:Cross-references: GB:M55661; NID:g145507; PIDN:AAC41416.1; PID:g145510  
A:Experimental source: enterotoxigenic strain, CFA/I-ST plasmid NTP113  
A:Note: sequence extracted from NCBI backbone (NCBIN:108960, NCBI:P:108971)  
C:Genetics:  
A:Gene: cfaC  
A:Genome: plasmid

Query Match 7.5%; Score 91.5; DB 2; Length 869;  
Best Local Similarity 22.3%; Pred. No. 11;  
Matches 41; Conservative 32; Mismatches 46; Indels 65; Gaps 13;

Qy 75 NVEGSY-----GVRQSF-----GKNQNPAGWTGNPNHVYK-----IEWLN 110  
Db 432 NAKGNVQSISYTDGFSLSFYHNDKRVDCGRNIN-AGWSG-CYESYSASLSIPLLGWTS 488  
Qy 111 GLUSYVGDFWN-----RSHLIADSLGGDALRVNAVVTGTRTONV-GGRDQKGMRYTEQRAQEW 166  
Db 489 TLGYSPTYSESIVYKSHILSE-----YGFYNQNIYKGRTOR--WOLTSSTSLKW 534  
Qy 167 LEARNRGYLYVEVAPYINAD--ELIPR-----AVVYSMQSSDNTINEKVLVYNTANGYTIN 220  
Db 535 MD-----YNFMPAIGLYNSEQRQLTKGQYISVTTIRASRENSLN-----TGYSYN 580  
Qy 221 YHNG 224  
Db 581 YSRG 584

RESULT 10  
T30295  
P-glycoprotein - Trypanosoma cruzi  
C:Species: Trypanosoma cruzi  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T30295  
R:Dallagiovanna, B.; Gamarro, F.; Castanys, S.  
Mol. Biochem. Parasitol. 75, 145-157, 1996  
A:Title: Molecular characterization of a P-glycoprotein-related tcpgp2 gene in Trypanoso  
A:Reference number: Z11742; MUID:96258555; PMID:8992313  
A:Accession: T30295  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1534 <DAL>  
A:Cross-references: EMBL:Z49222; NID:g1150649; PIDN:CAA89197.1

C:Genetics:  
A:Note: TCPGP2  
C:Superfamily: human multidrug resistance protein cMOAT2; ATP-binding cassette homology

Query Match 7.5%; Score 91.5; DB 2; Length 1534;  
Best Local Similarity 25.4%; Pred. No. 23;  
Matches 45; Conservative 23; Mismatches 72; Indels 37; Gaps 8;

Qy 59 DELGRTRTARGTLTYANVEGSGYVRSQSPGKNP---ACW-----TCNP---NH 101  
Db 1224 DEL-RRRTTRAADVTVTFGEASAR---SKGQHPETTAGWLEFREVEMRYAGLPLVLDR 1279  
Qy 102 VKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNAVVTGTRTONVGGRDQKGMRYTEQ 161  
Db 1280 VSPRIEPRQKGVVG---RTSGKSTLAATLMRWVEICGRRHVRVGAADRLYGLRELQ 1335  
Qy 162 R-----AQEWLEARNRGYLYVEVAPYINADDELIPRAVVVSMQSS--DNTINE 206  
Db 1336 QFSMIPQDPVLFDTGTVRSNLDPLDSTPAEVMRALELVGMRELLESGGIDGRVQE 1392

RESULT 11  
AB1481  
conserved hypothetical protein, probable lipoprotein lin0385 [imported] - Listeria innoc  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AB1481  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.  
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat  
ok, C.; Schluster, T.; Simoes, N.; Tiertez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,  
A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1481  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-386 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CAC95618.1; PID:g16412814; GSPDB:GN00178  
A:Experimental source: strain Clp11262  
C:Genetics:  
A:Gene: lin0385

Query Match 7.5%; Score 91; DB 2; Length 386;  
Best Local Similarity 21.2%; Pred. No. 4.2;  
Matches 55; Conservative 41; Mismatches 81; Indels 82; Gaps 15;

Qy 28 NDSPNYK-----TLGTSQITPA-----LFPKAGDILY-----SKLDEL----- 61  
Db 113 NDSNFDKLDNATGLDASLDTLAKAKETKAKTSEVLDKAVDNYKYVTDQVDFVQAOTE 172  
Qy 62 -----GRTRTARGT-----LTYANVEGSGYVRSQSF-----KNQNPAGWTG 97  
Db 173 IFTNAVKSQDIEXAKATYVSPRLNYERIE---PIAESFGDLDPKIDARINDVNEADWTG 229  
Qy 98 NPNHVYKIEW-----LNGLSYVGDFWNRSHLIADSLGGDALRVNAVVTGTRTONVGGRDQ- 152  
Db 230 --FVIERALWEKSKLDGMD-----VYADKLJLTDALQA-----EVGNLKEPKP 273  
Qy 153 --KGMRYTEQRAQEWLEARNRGYLYVEVAPYINAD-----ELIPRAVVVSMQSSDNTINE 206  
Db 274 MVAGAMELLNEAATTKITGEEAYSHTDLDL--NANVEGSKVYVQAIIIPALNAQDKDLAD 332  
Qy 207 KV-LVYNTANGYTYNHNG 224  
Db 333 QIDAAPNMDTLANYKNG 351

RESULT 12  
S27545  
pullulanase - Thermoaerobacterium thermosulfurigenes  
C:Species: Thermoaerobacterium thermosulfurigenes

C>Date: 09-Jun-1994 #sequence\_revision 26-May-1995 #text\_change 26-May-1995

C/Accession: S27545

R:Burchhardt, G.; Haeckel, K.; Sprinat, A.; Antranikian, G.; Bahl, H.

submitted to the EMBL Data Library, March 1992

A:Description: Nucleotide sequence of the pullulanase gene from Clostridium thermosulfur

A:Reference number: S27544

A:Accession: S27545

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1203 <BUR>

A:Cross-references: EMBL:M57692

A>Note: the source is given as Clostridium thermosulfurogenes

Query Match 7.5%; Score 91; DB 2; Length 1203;

Best Local Similarity 21.6%; Pred. No. 19;

Matches 63; Conservative 36; Mismatches 86; Indels 106; Gaps 18;

Qy 4 QVSDVNLGASKYLNALAWTNDSPNYKTLGT-----SQITPALFPKAG- 51

Db 931 QAPSNVVTSGKV---DLSWLSQDGTGATGNIYRSSVEGGLYKIASNVTTTFEDANV 987

Qy 52 ----DILY--SKLDELGRTTARTGLTYANVEGSGVRSFGKQNPAGWTGN-----PN 100

Db 988 TNGLYVYVAISAIDELG---NESGISNDVAVPAY-----PIGWGNLTQVSDN 1033

Qy 101 HV-----KYKIEWLNGLS-----YVGFNWRSHLIADSLGGDALR 135

Db 1034 HIIGVDKPTEDIYAEVWADGLTNGTGGPNMIAQLGYKYV-----SGTVYDVSYGSV-- 1085

Qy 136 VNAVGTGRTQ-----NVGGRDQ-----KGG-----MRYTEQRAQWLEANR 171

Db 1086 YNSVYGVDDSGFTWNAQYGVGDGNDQYKASFTPDKIGOWEYLMRFSNDQGDWITTT 1145

Qy 172 DGYLYVEVAPYNADELI-PRVVVSMQSSDNTINEKVLVNTANGYTYNY 221

Db 1146 ---LSFYVVP---SDDLKPTAPYLNQPGTES---SRVSLTWNPTDVGNY 1188

RESULT 13

S58857

botulinum neurotoxin type B hemagglutinin component, 33K - Clostridium botulinum (strain

N;Alternate names: protein HA-33

C/Species: Clostridium botulinum

A:Variety: strain Eklund 17B

C/Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 15-Oct-1999

C/Accession: S58857

R:Eaet, A.K.; Stacey, J.M.; Collins, M.D.

Syst. Appl. Microbiol. 17, 306-312, 1994

A>Title: Cloning and sequencing of a hemagglutinin component of the botulinum neurotoxin

A:Reference number: S58855

A:Accession: S58857

A:Molecule type: DNA

A:Residues: 1-292 <EAS>

A:Cross-references: EMBL:X79103; NID:g870932; PIDN:CAAS55714.1; PID:g870935

A:Experimental source: strain Eklund 17B

C/Keywords: hemagglutinin; neurotoxin

Query Match 7.4%; Score 90.5; DB 2; Length 292;

Best Local Similarity 22.6%; Pred. No. 3.2;

Matches 48; Conservative 28; Mismatches 67; Indels 69; Gaps 12;

Qy 74 ANVEGSGY-VRSFGKQNPAGWTGNPNHVKY-----KIEWLNGLSY-----V 115

Db 88 SNADNQYWLKLLDKIGNSFIASYNPNLVYADTVARNLKLSTLNNSSVIKFIIDYMI 147

Qy 116 GDFWN-----RSHLIADSLGGDALRVNAVVTGRTQNVGGRDQGGWRYTEQRAQ-- 164

Db 148 SDFNFTCKISPILDSSKVQVAMTDLNLYTWDY-----GRNQWTKYNKEKSAYQ 202

Qy 165 -----EWLEAN-----RDGLYVEVAPYNADELI-PRVVVSMQSSDN 202

Db 203 PFNTILNGVLTVISSNGNTVRVSSIAQNNDQAYWLINPVSNAYETY---TITNLHDTTK 259

Qy 203 TINEKVLVYN--TANGYTI---NVHNGTPTQK 229

Db 260 ALD-----LYNSQTANGTTIQVFNYH-GDDNQK 286

RESULT 14

F89887

Alpha-Hemolysin precursor [imported] - Staphylococcus aureus (strain N315)

C/Species: Staphylococcus aureus

C/Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 01-Mar-2002

C/Accession: F89887

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguci,

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C;Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A>Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F89887

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-319 <KUR>

A:Cross-references: GB:BA000018; PID:g13700962; PIDN:BAB42258.1; GSPDB:GN00149

A:Experimental source: strain N315

C/Genetics:

A:Gene: SAI007

C:Superfamily: leukocidin

Query Match 7.4%; Score 90; DB 2; Length 319;

Best Local Similarity 18.9%; Pred. No. 4;

Matches 46; Conservative 43; Mismatches 88; Indels 66; Gaps 10;

Qy 29 DSPNYKTLGTSQITPALFPKAGD-----ILYSKLDLGTGRTTARTGLTYAN 75

Db 28 DSDINIKT-CTTDIGSNTTVKTDGLVTDYDKENGHKKVYFSDIDDKHKKLLVIRTKGT 86

Qy 76 VEGSYGVRSFGKQNPAGWTGNPNHVKYIWL-NGLSVYGVDPWNSHL-----IA 126

Db 87 IAGQYRVYSEEGANKSGLAW---PSAPKVLQPLPDNEVAQISDYPRNSIDTKEYMSTLT 143

Qy 127 DSLGGDALRVNAVVTGRTQNVGG-----RDQKGMK--Y 158

Db 144 YGFNGN-----VTGDDTGKIGLGANVSIGHTLKVQPDFKTLILSPDCKVGVKVI 197

Qy 159 TEQRAQEWLEANDGYLYVEVAPYNADELI-PRVVVSMQSSDNTINEKVLVNTANGYT 218

Db 198 NNMVNQWGPYDRDSW-----NPVY-GNQLFMKTRNGSKAAENFLDPNKASLLSGFS 251

Qy 219 INY 221

Db 252 PDF 254

RESULT 15

F97223

uncharacterized conserved protein, YOME B. subtilis ortholog CAC2630 [imported] - Clostri-

C/Species: Clostridium acetobutylicum

C/Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C/Accession: F97223

R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: F97223

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-347 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK80577.1; PID:g15025656; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C/Genetics:

A:Gene: CAC2630

Query Match 7.3%; Score 89.5; DB 2; Length 347;

[illegible]

Search completed: January 5, 2004, 18:42:52  
Job time : 27.5691 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	365	29.9	327	1	DRN1_STREQ	P26295 streptococc
2	122.5	10.0	274	1	NUCE_STRPN	Q03158 streptococc
3	99.5	8.2	1279	1	APU_THESA	P36905 t amyltopull
4	97	8.0	1861	1	APU_THETP	P38536 t amyltopull
5	94	7.7	319	1	HLA_STRAU	P09616 staphylococ
6	91.5	7.5	869	1	CFAC_ECOLI	P25733 escherichia
7	90	7.4	341	1	CMPL_PROPR	Q52581 photobacter
8	88.5	7.3	1476	1	GTEP_STRMU	P08987 streptococc
9	88.5	7.3	2208	1	POLN_MANCV	Q69014 manchester
10	88	7.2	802	1	PAC_ARTVI	P31956 arthroabacte
11	87.5	7.2	1047	1	POL_SIVAI	P27973 simian immu
12	87	7.1	420	1	DHE3_PVRFU	P80319 pyrococcus
13	87	7.1	420	1	DHE3_PVRHO	O52310 pyrococcus
14	86.5	7.1	1116	1	RPOB_HETCA	P36440 heterosigma
15	86.5	7.1	1709	1	CHD1_HUMAN	O14646 homo sapien
16	86	7.0	296	1	NIFD_NOSCO	P52337 nostoc comm
17	86	7.0	1182	1	RPOB_STRAU	P27768 staphylococ
18	85	7.0	420	1	DHE3_PVRAB	Q47950 pyrococcus
19	85	7.0	1165	1	RPC2_SCHPO	Q10233 schizosacch
20	84	6.9	362	1	OMPF_ECOLI	P02931 escherichia
21	84	6.9	1102	1	RPOB_SVNY3	P77965 synechocvat
22	83.5	6.8	3093	1	POLG_BSTV1	Q65730 b genome po
23	83	6.8	227	1	XYNI_HUMIN	P55334 hemicholia in
24	83	6.8	398	1	KBL_ECOLI	P07912 escherichia
25	83	6.8	398	1	KBL_SALTY	P37419 salmonella
26	82.5	6.8	480	1	NIFD_CYAA5	Q07642 cyanothece
27	82.5	6.8	1711	1	CHD1_MOUSE	P40201 mus musculu
28	82	6.7	392	1	PRTD_STRGR	P52321 streptomyce
29	81.5	6.7	1723	1	PM30_CHLPN	Q92812 chlamydia p
30	81	6.6	420	1	DHE3_PVREN	Q47951 pyrococcus
31	81	6.6	638	1	YQGS_BACSU	P54496 bacillus su
32	81	6.6	666	1	MUR2_ENTHR	P39046 enterococcu
33	81	6.6	1196	1	AMYE_PAEPO	P21543 paenibacill

162

RAQWLEARNRGVLYEYVAPTYNADELIPRAVVMQSSDNTINEKVLVYNTANGTYINY

221

KVLSHITKPNPVVYFSAIPEYQGAELARSVLVSALSDGVINETVRVENTADGFNINY

222

HNG 224

281

EKG 283

RESULT 2

NUCE STRPN

ID

NUCE STRPN

STANDARD;

PRT;

274

AA.

AC

Q03158;

DT

01-JUN-1994 (Rel. 29, Created)

DT

01-JUN-1994 (Rel. 29, Last sequence update)

DT

15-SEP-2003 (Rel. 42, Last annotation update)

DE

DNA-entry nuclease (Competence-specific nuclease) (EC 3.1.30.-).

OS

Streptococcus pneumoniae, and

OS

Streptococcus pneumoniae (strain ATCC BAA-255 / R6).

OC

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC

Streptococcus

OX

NCBI\_taxid=1313, 171101;

RN

[1]

RP

SEQUENCE FROM N.A.

RC

STRAIN=470;

RC

MEDLINE=90294291; PubMed=2359120;

RA

Puyet A., Greenberg B., Lacks S.A.;

RT

"Genetic and structural characterization of enda. A membrane-bound

RT

nuclease required for transformation of Streptococcus pneumoniae";

RL

J. Mol. Biol. 213:727-738(1990).

RN

[2]

RP

SEQUENCE FROM N.A.

RC

STRAIN=ATCC BAA-334 / TIGR4;

RC

MEDLINE=21357209; PubMed=11463916;

RA

Tetzelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,

RA

Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,

RA

Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,

RA

Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,

RA

Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,

RA

McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,

RA

Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,

RA

Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;

RT

"Complete genome sequence of a virulent isolate of Streptococcus

RT

pneumoniae";

RL

Science 293:498-506(2001).

RN

[3]

RP

SEQUENCE FROM N.A.

RC

STRAIN=ATCC BAA-255 / R6;

RC

MEDLINE=21429245; PubMed=11544234;

RA

Hoskins J., Alborn W.B. Jr., Arnold J., Blaszcak L.C., Burgett S.,

RA

DeHoff B.S., Estrom S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,

RA

Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,

RA

LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushima P.,

RA

McNair S.M., McHenry M., McLeaster K., Mundy C.W., Nicas T.I.,

RA

Norris F.H., O'Garra M., Peery R.B., Robertson G.T., Rockey P.,

RA

Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,

RA

Zook C.A., Baltz R.H., Jaskunas S.R., Rostek P.R. Jr., Skatrud P.L.,

RA

Glass J.I.;

RT

"Genome of the bacterium Streptococcus pneumoniae strain R6";

RL

J. Bacteriol. 183:5709-5717(2001).

CC

-I- FUNCTION: By degrading DNA that enters the cell, plays a role in

CC

the competence of cells to be transformed.

CC

-I- SUBCELLULAR LOCATION: Membrane-bound.

CC

-I- SIMILARITY: Belongs to the DNA/RNA non-specific endonuclease

CC

family.

CC

-----

CC

This SWISS-PROT entry is copyright. It is produced through a collabora-

CC

tion between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC

the European Bioinformatics Institute. There are no restrictions on its

CC

use by non-profit institutions as long as its content is in no way

CC

modified and this statement is not removed. Usage by and for commercial

CC

-----

CC

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

-----

DR

EMBL; X54225; CAA38134.1; -.

DR

EMBL; AE007486; AAK76031.1; -.

DR

EMBL; AE008543; AAL00582.1; -.

DR

PIR; A99094; A99094.

DR

PIR; F95229; F95229.

DR

PIR; S10641; S10641.

DR

TIGR; SP1964; -.

DR

InterPro; IPR001604; Endonuclease.

DR

Pfam; PF01223; Endonuclease; 1.

DR

PROSITE; PS01070; NUCLEASE NON SPEC; 1.

KW

Competence; Hydrolase; Nuclease; Endonuclease; Signal-anchor;

KW

Transmembrane; Complete proteome.

FT

TRANSMEM 8 25 SIGNAL-ANCHOR (POTENTIAL).

FT

ACT SITE 160 160 BY SIMILARITY.

SQ

SEQUENCE 274 AA; 29890 MW; 59B2243F0150CD98 CRC64;

Query Match

10.0%; Score 122.5; DB 1; Length 274;

Best Local Similarity

22.3%; Pred. No. 0.0011;

Matches

57; Conservative 34; Mismatches 100; Indels 65; Gaps 11;

Qy

3

TQVSNVDWLNDGASKYLEALAWTND-----SPNYKTLGTSGI 42

Db

50

SQALAESVLTDAVKSQIKGLEWNGSGAFIVNGNKTNLDAKVSCKPYADNKTGKTV 109

Qy

43

TPALFPKAGDILYSKLDE--LGRTRTARGLTYANVEGSGYGRQSGKQNPAGWTGNPN 100

Db

110

-----PTVALLSKATROYKVKETGNGSTSW-----TPGW----- 142

Qy

101

HVKYKIEWLNGLSYVGDFWNRSHLIADSL-----GGDALR---VNAVTRTRTONVGRDQ 152

Db

143

---HQVKNLKG-SYTHAV--DRGHLGLYALIGGLDGFDASTSNPKNIATVATAMANOQAQY 197

Qy

153

KGMRYTEORAQEWLEARDRGHLYEVAPIYNADE--LIPRAVVMQSSDNTINEKVLVY 211

Db

198

STQNYVESKVRKALDQNK--RVRYRVTLTYASNEDLVPSASQIEAKSSDGELEFNVLP 255

Qy

212

NTANGTYINVHGTPT 227

Db

256

NVQKGLQDVRTGEVT 271

RESULT 3

APU\_THESA

ID

APU\_THESA

STANDARD;

PRT;

1279

AA.

AC

P36905;

DT

01-JUN-1994 (Rel. 29, Created)

DT

01-FEB-1995 (Rel. 31, Last sequence update)

DT

28-FEB-2003 (Rel. 41, Last annotation update)

DE

Amylopullulanase precursor (Alpha-amylase/pullulanase) [includes:

DE

Alpha-amylase (EC 3.2.1.1) (1,4-alpha-D-glucan glucanohydrolase);

DE

Pullulanase (EC 3.2.1.41) (1,4-alpha-D-glucan glucanohydrolase)

DE

(Alpha-dextrin endo-1,6-alpha-glucosidase)].

GN

APU.

OS

Thermoanaerobacter saccharolyticum.

OS

Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;

OC

Thermoanaerobacteriaceae; Thermoanaerobacterium.

OX

NCBI\_TaxID=28896;

RN

[1]

RP

SEQUENCE FROM N.A.

RC

STRAIN=B6A-RI;

RC

MEDLINE=94161525; PubMed=8117096;

RA

Ramesh M.V., Podkovyrov S.M., Lowe S.E., Zeikus J.G.;

RT

"Cloning and sequencing of the Thermoanaerobacterium saccharolyticum

RT

B6A-RI apu gene and purification and characterization of the

RT

amylopullulanase from Escherichia coli";

RL

Appl. Environ. Microbiol. 60:94-101(1994).

RN

[2]

RP

IDENTIFICATION OF PROBABLE VECTOR CONTAMINATION.

RA

Robison K.;

RL

Unpublished observations (NOV-1994).



CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic  
 CC linkages in pullulan and in amylopectin and glycogen, and the  
 CC alpha- and beta-limit dextrins of amylopectin and glycogen.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.  
 CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN BY THE PRESENCE OF  
 CC AN EXTRA C-TERMINAL SEGMENT OF 9 RESIDUES THAT SEEMS TO ORIGINATE  
 CC FROM A PUC-TYPE VECTOR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L07762; AAA19800.1; ALT S80.  
 CC HSP; Q08751; 1BVZ.  
 CC InterPro; IPR006589; Alp\_ami1\_cat sub.  
 CC InterPro; IPR006048; Alpha\_ami1\_C.  
 CC InterPro; IPR006047; Alpha\_ami1\_cat.  
 CC InterPro; IPR003961; FN\_III.  
 CC InterPro; IPR004185; Glyco\_hydro\_13Ig.  
 CC Pfam; PF00128; alpha-amy1ase; 1.  
 CC Pfam; PF02806; alpha-amy1ase\_C; 1.  
 CC Pfam; PF02903; alpha-amy1ase\_N; 1.  
 CC Pfam; PF00041; fn3; 2.  
 CC SMART; SM00642; Amy; 1.  
 CC SMART; SM00632; Amy\_C; 1.  
 CC SMART; SM00060; FN3; 2.  
 CC Hydroxylase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;  
 KW Multifunctional enzyme.  
 FT SIGNAL 1 35 POTENTIAL.  
 FT CHAIN 36 1279 AMYLOPULLULANASE.  
 FT DOMAIN 929 1017 FIBRONECTIN TYPE-III 1.  
 FT DOMAIN 1156 1248 FIBRONECTIN TYPE-III 2.  
 FT ACT\_SITE 629 629 BY SIMILARITY.  
 FT ACT\_SITE 658 658 BY SIMILARITY.  
 FT ACT\_SITE 735 735 BY SIMILARITY.  
 SQ SEQUENCE 1279 AA; 142430 MW; 095CCBCA391624DD CRC64;  
 Query Match 8.2%; Score 99.5; DB 1; Length 1279;  
 Best Local Similarity 22.6%; Pred. No. 0.88;  
 Matches 69; Conservative 33; Mismatches 80; Indels 123; Gaps 19;  
 QY 4 QVNDVVLNDGASKYLNEALWTFNDSPNYYKLTGTSQITPALFPKAGD----- 52  
 Db 932 QVPSNVVATSGNGKV---DLSSQSDGATGY-NYRSSVEGGLYEKIASNVVTGTTFDTN 987  
 QY 53 -----ILYSKLDLGL-RTTRAGTLTYANVEGSGYGVRSQKQKQNPAGWTGN----- 98  
 Db 988 VTNGLKYVVAISAVALDELNESEMSIDTVAYP-----AY-----PIGWNLTVQVV 1032  
 QY 99 PNHV-----KYKIEWLNGLS-----YVGDFWNRSHLIADSLGDA 133  
 Db 1033 DNHVISVSNPTEDIYAEVWADGLTNSTGQGNMIAQLGYKYVGGTYN-----DSVYGSV 1086  
 QY 134 LRNVAVTGTRTQ-----NVGGRDQKGG-----MRYTEQRAQEWLEAN 170  
 Db 1087 --YNSVGVVDDSDFTWNAQYVGDIGNDQYKASLHLINRSMGYLMRPSDNQGSQWTTTD 1144  
 QY 171 RDGLYLYEVAPIYNADELI-PRAVV-----VSMOSSDNTINEKVLVYN-----TAN 215  
 Db 1145 T---LSFYVVP---SDDLKPTAPILNQGVSSRSVLTWSPSTDN--VGIVNVEIYRSD 1196  
 QY 216 GYTIN 220  
 Db 1197 GGTFN 1201

RESULT 4  
 APU THETU  
 ID APU THETU STANDARD; PRT; 1861 AA.  
 AC P38536;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Amylopullulanase precursor (Alpha-amy1ase/pullulanase) (Pullulanase  
 DE type II) [Includes: Alpha-amy1ase (EC 3.2.1.1) (1,4-alpha-D-glucan  
 DE glucanohydrolase); Pullulanase (EC 3.2.1.41) (1,4-alpha-D-glucan  
 DE glucanohydrolase) (Alpha-dextrin endo-1,6-alpha-glucosidase)].  
 GN AMYB.  
 OS Thermoanaerobacter thermosulfurogenes (Clostridium  
 OS thermosulfurogenes).  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacterium.  
 OX NCBI\_TaxID=33950;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 3896 / EMI;  
 RC MEDLINE=94252998; PubMed=8195085;  
 RX Matuschek M., Burchhardt G., Sahm K., Bahl H.;  
 RA "Pullulanase of Thermoanaerobacterium thermosulfurogenes EMI  
 RT (Clostridium thermosulfurogenes): molecular analysis of the gene,  
 RT composite structure of the enzyme, and a common model for its  
 RT attachment to the cell surface.";  
 RL J. Bacteriol. 176:3295-3302(1994).  
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-alpha-glucosidic  
 CC linkages in oligosaccharides and polysaccharides.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of (1->6)-alpha-D-glucosidic  
 CC linkages in pullulan and in amylopectin and glycogen, and the  
 CC alpha- and beta-limit dextrins of amylopectin and glycogen.  
 CC -1- SUBCELLULAR LOCATION: CELL-BOUND. IT C-TERMINUS MAY SERVE AS AN  
 CC S-LAYER ANCHOR.  
 CC -1- PTM: GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO  
 CC KNOWN AS THE ALPHA-AMYLASE FAMILY.  
 CC -1- SIMILARITY: Contains 2 fibronectin type III domains.  
 CC -1- SIMILARITY: Contains 3 S-layer homology (SLH) domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M57692; AAB00841.1; -  
 CC HSP; Q08751; 1BVZ.  
 CC InterPro; IPR006589; Alp\_ami1\_cat sub.  
 CC InterPro; IPR006048; Alpha\_ami1\_C.  
 CC InterPro; IPR006047; Alpha\_ami1\_cat.  
 CC InterPro; IPR003961; FN\_III.  
 CC InterPro; IPR004185; Glyco\_hydro\_13Ig.  
 CC InterPro; IPR004193; Glyco\_hydro\_13N.  
 CC InterPro; IPR001119; SLH  
 CC Pfam; PF00128; alpha-amy1ase; 1.  
 CC Pfam; PF02806; alpha-amy1ase\_C; 1.  
 CC Pfam; PF02903; alpha-amy1ase\_N; 1.  
 CC Pfam; PF00041; fn3; 2.  
 CC Pfam; PF02922; isoamy1ase\_N; 1.  
 CC Pfam; PF00395; SLH; 3.  
 CC SMART; SM00642; Amy; 1.  
 CC SMART; SM00632; Amy\_C; 1.  
 CC SMART; SM00060; FN3; 2.  
 CC PROSITE; PS01072; SLH DOMAIN; 3.  
 KW Hydroxylase; Glycosidase; Carbohydrate metabolism; Signal; Repeat;  
 KW Multifunctional enzyme; Glycoprotein.  
 FT SIGNAL 1 35 POTENTIAL.  
 FT CHAIN 36 1861 AMYLOPULLULANASE.  
 FT DOMAIN 928 1018 FIBRONECTIN TYPE-III 1.

FT	DOMAIN	1157	1248	FIBRONECTIN TYPE-III 2.
FT	ACT_SITE	628	628	BY SIMILARITY.
FT	ACT_SITE	657	657	BY SIMILARITY.
FT	ACT_SITE	734	734	BY SIMILARITY.
FT	DOMAIN	1681	1739	SLH 1.
FT	DOMAIN	1740	1803	SLH 2.
FT	DOMAIN	1804	1861	SLH 3.
FT	CONFLICT	1734	1734	D->E (IN REF. 1; AAB00841).
FT	SEQUENCE	1861	AA; 206104 MW; 06C23070E45B574 CRC64;	
Query Match				
Best Local Similarity 21.4%; Pred. No. 2.4;				
Matches 70; Conservative 35; Mismatches 82; Indels 140; Gaps 20;				
Qy	4	QVSDVNDGASKYLNEALAWTFDPSNYKILGT-----SQITPALFPKAG- 51		
Db	931	QAPSNVVTSGKGV---DLSWLQSDGATGNYIRSVSEGLYEKIASNVTETTFDANV 987		
Qy	52	----DILY--SKDELGRTRTARGLTYANVEGSGYGVRFQGNQNPAGWTGN-----PN 100		
Db	988	TNGLKYVTAISAIDELG---NESGISNDVAVPAY-----PIGWGLTQVSDN 1033		
Qy	101	HV-----KYKIEWLNGLS-----YVGFDFNRSHLIADSLGGDALR 135		
Db	1034	HIIGVDKPTEDIYAEVWADGLTNSTGQPNMIAQLGYKV-----SGTVDSYVGSV-- 1085		
Qy	136	VNAVGTGRTQ-----NVGRDQ-----KGG-----MRYTEORAEWLEANR 171		
Db	1086	YNSVYGVDDSGFTWVNAQYVGDIGNPDQYKASFTPDKIGOWEYLMRFSNDQGDWITST 1145		
Qy	172	DGVLVYEVAPIYNADELI--PRAVVVSM-----QSSDN----- 202		
Db	1146	---LSFYVP---SDDLKPTAPYLNQPGTSSRSVSLTWPSTDNVGIYDIYRSDDGT 1199		
Qy	203	-----TINEKLVY---NTANGYTYNY 221		
Db	1200	FNKIATSVNEVYVYDTSVINGVTYNY 1226		
RESULT 5				
ID	HLA_STAAU	STANDARD;	PRT;	319 AA.
AC	P09616;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Alpha-hemolysin precursor (Alpha-toxin) (Alpha-HL).			
GN	HLY OR HLA.			
OS	Staphylococcus aureus.			
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.			
OX	NCBI_TaxID=1280;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 27-44.			
RC	STRAIN=Wood 46;			
RX	MEDLINE=85053471; PubMed=6500704;			
RA	Gray G.S., Kehoe M.;			
RT	"Primary sequence of the alpha-toxin gene from Staphylococcus aureus			
RT	wood 46.";			
RL	Infect. Immun. 46:615-618(1984).			
RN	[2]			
RP	REVIEWS, SEQUENCE FROM N.A.			
RC	STRAIN=Wood 46;			
RA	Hedengrahn G.;			
RL	Submitted (OCT-1992) to the EMBL/GenBank/DBSJ databases.			
RN	[3]			
RP	SEQUENCE OF 27-319 FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=Wood 46;			
RX	MEDLINE=92268149; PubMed=1587866;			
RA	Walker B., Krishnasastri M., Zorn L., Kasianowicz J., Bayley H.;			
RT	"Functional expression of the alpha-hemolysin of Staphylococcus			
RT	aureus in intact Escherichia coli and in cell lysates. Deletion of			
RT	five C-terminal amino acids selectively impairs hemolytic			
RT	activity.";			

RL	J. Biol. Chem. 267:10902-10909(1992).			
RN	[4]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RC	STRAIN=Wood 45 / ATCC 10832;			
RX	MEDLINE=97102581; PubMed=8943190;			
RA	Song L., Hobaugh M.R., Shustak C., Cheley S., Bayley H., Gouaux J.E.;			
RT	"Structure of staphylococcal alpha-hemolysin, a heptameric			
RT	transmembrane pore.";			
RL	Science 274:1859-1866(1996).			
RN	[5]			
RP	MUTAGENESIS.			
RX	MEDLINE=93016135; PubMed=1400487;			
RA	Walker B., Krishnasastri M., Zorn L., Bayley H.;			
RT	"Assembly of the oligomeric membrane pore formed by Staphylococcal			
RT	alpha-hemolysin examined by truncation mutagenesis.";			
RL	J. Biol. Chem. 267:21782-21786(1992).			
RN	[6]			
RP	MUTAGENESIS OF HISTIDINE RESIDUES.			
RX	MEDLINE=94222552; PubMed=8168947;			
RA	Menzies B.E., Kernodle D.S.;			
RT	"Site-directed mutagenesis of the alpha-toxin gene of Staphylococcus			
RT	aureus: role of histidines in toxin activity in vitro and in a murine			
RT	model.";			
RL	Infect. Immun. 62:1843-1847(1994).			
RN	[7]			
RP	MUTAGENESIS OF HISTIDINE RESIDUES.			
RX	MEDLINE=94245329; PubMed=8188346;			
RA	Jursch R., Hildebrand A., Hobom G., Tranum-Jensen J., Ward R.;			
RA	Kehoe M., Bhakdi S.;			
RT	"Histidine residues near the N-terminus of staphylococcal alpha-toxin			
RT	as reporters of regions that are critical for oligomerization and			
RT	pore formation.";			
RL	Infect. Immun. 62:2249-2256(1994).			
RN	[8]			
RP	MUTAGENESIS.			
RX	MEDLINE=96032742; PubMed=7559447;			
RA	Walker B., Bayley H.;			
RT	"Key residues for membrane binding, oligomerization, and pore forming			
RT	activity of staphylococcal alpha-hemolysin identified by cysteine			
RT	scanning mutagenesis and targeted chemical modification.";			
RL	J. Biol. Chem. 270:23065-23071(1995).			
CC	-I- FUNCTION: ALPHA-TOXIN BINDS TO THE MEMBRANE OF EUKARYOTIC CELLS			
CC	RESULTING IN THE RELEASE OF LOW-MOLECULAR WEIGHT MOLECULES AND			
CC	LEADING TO AN EVENTUAL OSMOTIC LYSIS. HEPTAMER OLIGOMERIZATION			
CC	AND PORE FORMATION IS REQUIRED FOR LYTIC ACTIVITY.			
CC	-I- SUBUNIT: SELF-ASSEMBLES TO FORM FIRST, A NONLYTIC OLIGOMERIC			
CC	INTERMEDIATE, AND THEN, A MUSHROOM-SHAPED HOMOEPTAMER STRUCTURE			
CC	OF 100 ANGSTROMS IN LENGTH AND UP TO 100 ANGSTROMS IN DIAMETER.			
CC	-I- SUBCELLULAR LOCATION: SECRETED AS A MONOMER. AFTER OLIGOMERIZATION			
CC	AND PORE FORMATION, THE COMPLEX IS TRANSLOCATED ACROSS THE			
CC	BILAYER, PROBABLY VIA THE GLY-RICH DOMAIN OF EACH STRAND.			
CC	-I- DOMAIN: THE MUSHROOM-SHAPED HEPTAMER IS COMPOSED OF A CAP DOMAIN			
CC	(COMPRISING 7 BETA SANDWICHES AND THE AMINO LATCHES OF EACH			
CC	PROTOMER), 7 RIM REGIONS WHOSE PROTRUDING STRANDS MAY INTERACT			
CC	WITH THE MEMBRANE BILAYER, AND THE STEM DOMAIN (52 ANGSTROMS IN			
CC	LENGTH, 26 ANGSTROMS IN DIAMETER) WHICH FORMS THE TRANSMEMBRANE			
CC	PORE.			
CC	-I- SIMILARITY: BELONGS TO THE AEROLYSIN FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
CC	EMBL; X01645; CAA25801.1; -			
DR	EMBL; M90536; AAA26598.1; -			
DR	PIR; S69209; S69209.			
DR	PDB; 7AHL; 14-JAN-98.			
DR	InterPro; IPR005831; Aer_hem.			
DR	InterPro; IPR005830; Aer_hem_leuk.			



Best Local Similarity 22.3%; Pred. No. 2.8;  
Matches 41; Conservative 32; Mismatches 46; Indels 65; Gaps 13;

QY 75 NVEGSY-----GVROSF-----GKQNPAGTGNPNHVYK-----IWLNL 110  
DB 432 NAKGNYQSIVTDGFSLSFYHNDKRVDCGRNIN-AGWSG--CYESYSASISIPLLGWT 488  
QY 111 GLSYVGDFWN---RSHLIADSLGGDALRVNAVGTGRTQNV--GGDQKGMRYTEQRAQEW 166  
DB 489 TLGYSYDTSYSEVYSKSHLSE-----YGFYNQNIYKGTQR--WOLTSSTSLKW 534  
QY 167 LEANRDGGLYVEVAPIYNAD--ELIPR-----AVVVSQSSDNTINEKVLVNTANGYTIN 220  
DB 535 MD-----YNFMPAIGYINSEORQLTDKGYISVITRASRENSLN-----TGYSYN 580  
QY 221 YHNG 224  
DB 581 YSRG 584

## RESULT 7

OMPL\_PROPR STANDARD; PRT; 341 AA.  
AC Q52581;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Porin-like protein L precursor.  
GN OMPL.  
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Photobacterium.  
OX NCBI\_TaxID=74109;  
RN SEQUENCE FROM N.A.  
RP MEDLINE=96326353; PubMed=8759872;  
RA Welch T.J., Bartlett D.H.;  
RT "Isolation and characterization of the structural gene for Ompl, a  
RT pressure-regulated porin-like protein from the deep-sea bacterium  
RT Photobacterium species strain SS9."  
RL J. Bacteriol. 178:5027-5031(1996).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane  
CC (Potential).  
CC -!- INDUCTION: IN RESPONSE TO ELEVATED HYDROSTATIC PRESSURE.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U59311; AAB50064.1; -;  
DR InterPro; IPR003229; OMP 2.  
DR InterPro; IPR001702; Porin Gram-ve.  
DR Pfam; PF00267; Gram-ve\_porins; 1.  
DR ProDom; PD000608; OMP 2; 1.  
DR Transmembrane; Porin; Signal; Outer membrane.  
FT SIGNAL 1 21  
FT CHAIN 22 341 PORIN-LIKE PROTEIN L.  
SQ SEQUENCE 341 AA; 36672 MW; ED3D66A9F3DEB585 CRC64;

Query Match 7.4%; Score 90; DB 1; Length 341;  
Best Local Similarity 25.3%; Pred. No. 1.1;  
Matches 46; Conservative 11; Mismatches 65; Indels 60; Gaps 8;

QY 69 GTLTANVSGSYGVROSF-----GKQNPAGTGNPNHVYKIEWLNLGYVGDWNRSHL 124  
DB 113 GQLVYKADSGSLMTDFDIDMAYHGNAG-----NKLAAADRTNNLSYGSF----- 161  
QY 125 IADSLGGDALRVNAVGTGRTQNVGGDQKGMRYTEQRAQEWLE 168

DB 162 ---DLNGDNLTVKA-----NYVFGSDENEGYSAAMAYAMDMLGFGAGYGEQDQSSKN 213  
QY 169 ANRD-----GVLXYEVAPIYNADLIPRAVVVSQSSDNTINEKVLV-----YNTANG 216  
DB 214 GNEDKTKGQAFGAISYTIISDFYSG-----LYQDSRNTVVNDLIDESTGYEFAAA 264  
QY 217 YT 218  
DB 265 YT 266

## RESULT 8

GTFB\_STRMU STANDARD; PRT; 1476 AA.  
AC P08987; O69381; O69384; O69387; O69390; O69396;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)  
GN GTFB OR SMU 1004.  
OS Streptococcus mutans.  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1309;  
RN SEQUENCE FROM N.A.  
RP STRAIN=GS-5.  
RC MEDLINE=87308013; PubMed=3040685;  
RA Shiroza T., Ueda S., Kuramitsu H.K.;  
RT "Sequence analysis of the gtfB gene from Streptococcus mutans."  
RL J. Bacteriol. 169:4263-4270(1987).  
RN SEQUENCE FROM N.A.  
RP STRAIN=MT4239 / Serotype C, MT4245 / Serotype E, MT4251 / Serotype F,  
RC MT4467 / Serotype E, and MT8148 / Serotype C;  
RX MEDLINE=98231643; PubMed=9570124;  
RA Fujiwara T., Terao Y., Hoshino T., Kawabata S., Ooshima T., Sobue S.,  
RA Kimura S., Hamada S.;  
RT "Molecular analyses of glucosyltransferase genes among strains of  
RT Streptococcus mutans."  
RL FEMS Microbiol. Lett. 161:331-336(1998).  
RN SEQUENCE FROM N.A.  
RP STRAIN=UAI59 / ATCC 700610 / Serotype C;  
RX MEDLINE=2295063; PubMed=12397186;  
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;  
RT "Genome sequence of Streptococcus mutans UAI59, a cariogenic dental  
RT pathogen."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS. THAT ARE THOUGHT  
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
CC -!- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl} (N) = D-  
CC fructose + {(1,6)-alpha-D-glucosyl} (N+1).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- DISEASE: DENTAL CARIES.  
CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA  
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES). GTF-S SYNTHESIZES  
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH  
CC FORMS OF GLUCANS.  
CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
CC BINDING PROTEIN FROM S.MUTANS.  
CC -!- SIMILARITY: Contains 10 cell wall binding repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial





CC phosphomonoester.  
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + [DNA] (N).  
 CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE  
 CC DETERMINED.  
 CC -1- MISCELLANEOUS: THE 155 ISOLATE IS FROM A MONKEY IMPORTED FROM  
 CC KENYA.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M29875; AAA91906.1; --  
 CC HSSP; P03366; IHRH.  
 CC MEROPS; A02.003; --  
 CC InterPro; IPR001995; Aspartate rtrv.  
 CC InterPro; IPR001969; Aspartate site.  
 CC InterPro; IPR001037; Integrase\_C.  
 CC InterPro; IPR003308; Integrase\_Zn.  
 CC InterPro; IPR002156; RNaseH.  
 CC InterPro; IPR001584; Rve.  
 CC InterPro; IPR000477; RVTse.  
 CC Pfam; PF00552; Integrase; 1.  
 CC Pfam; PF02022; Integrase\_Zn; 1.  
 CC Pfam; PF00075; rnaseH; 1.  
 CC Pfam; PF00665; rve; 1.  
 CC Pfam; PF00077; rvp; 1.  
 CC Pfam; PF00078; rvt; 1.  
 CC PROSITE; PS00141; ASP\_PROTEASE; 1.  
 CC PROSITE; PS00175; ASP\_PROT\_RETROV; 1.  
 CC AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;  
 CC Nuclease; Transferase; RNA-directed DNA polymerase.  
 CC ACT SITE 113 113 BY SIMILARITY.  
 CC SQ SEQUENCE 1047 AA; 118871 MW; A38DDDA39269BE5 CRC64;  
 CC -----  
 CC Query Match 7.2%; Score 87.5; DB 1; Length 1047;  
 CC Best Local Similarity 28.4%; Pred. No. 7.9;  
 CC Matches 40; Conservative 16; Mismatches 48; Indels 37; Gaps 9;  
 CC -----  
 CC QY 51 GDILYS-KLDELGRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNP----- 99  
 CC DB 303 GDAYSPLDPEFRKYTA-FTIPTVNEGP-GIRYQF--NCLPGWKGSPITFQNTASKI 359  
 CC QY 100 -NHVKYKIEWLNGLSYGVDFWNRSHLIADSLGSDALRVNAVTRTQNVGGRDQKGMRY 158  
 CC DB 359 LEEIKKELKQLTIYQYMDLLWVGSG--EEGPKHDQL-----VQLRLNLQEWGLET 407  
 CC QY 159 TEQRAQ-----EWLEANDRGY 174  
 CC DB 408 PEKKVQREPPPEWM-----GY 423  
 CC -----  
 CC RESULT 12  
 CC DHE3\_PYRFU  
 CC ID DHE3\_PYRFU STANDARD; PRT; 420 AA.  
 CC AC P80319;  
 CC DT 01-JUN-1994 (Rel. 29, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 CC DE Glutamate dehydrogenase (EC 1.4.1.3) (GDH).  
 CC GN GDHA OR GDH OR PFI602.  
 CC OS Pyrococcus furiosus.  
 CC OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 CC CC Pyrococcus.  
 CC OX NCBI\_TaxID=2261;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.

RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RX MEDLINE=94010338; PubMed=8406037;  
 RA Eggen R.L.L., Geerling A.C.M., Waldoetter K., Antranikian G.,  
 RA de Vos W.M.;  
 RT "The glutamate dehydrogenase-encoding gene of the hyperthermophilic  
 RT archaeon Pyrococcus furiosus: sequence, transcription and analysis of  
 RT the deduced amino acid sequence.";  
 RL Gene 132:143-148(1993).  
 RN [2]  
 RP SEQUENCE.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RX MEDLINE=94338538; PubMed=8060497;  
 RA Maras B., Valiente S., Chiaraluce R., Consalvi V., Politi L.,  
 RA de Rosa M., Bessa F., Scandurra R., Barra D.;  
 RT "The amino acid sequence of glutamate dehydrogenase from Pyrococcus  
 RT furiosus, a hyperthermophilic archaeobacterium.";  
 RL J. Protein Chem. 13:253-259(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RX Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RX MEDLINE=96164432; PubMed=8591026;  
 RA Yip K.S.P., Stillman T.J., Britton K.L., Artymuk P.J., Baker P.J.,  
 RA Sedelnikova S.E., Engel P.C., Pasquo A., Chiaraluce R., Consalvi V.,  
 RA Scandurra R., Rice D.W.;  
 RT "The structure of Pyrococcus furiosus glutamate dehydrogenase reveals  
 RT a key role for ion-pair networks in maintaining enzyme stability at  
 RT extreme temperatures.";  
 RL Structure 3:1147-1158(1995).  
 CC -1- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NAD(P)(+) = 2-  
 CC oxoglutarate + NH(3) + NAD(P)H.  
 CC -1- SUBUNIT: Homohexamer.  
 CC -1- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M97860; AAA83390.1; --  
 CC EMBL; AE010260; AAL81726.1; --  
 CC PIR; T46971; JN0854.  
 CC PDB; 1GTW; 11-JAN-97.  
 CC InterPro; IPR006095; GLFV dehydrog.  
 CC InterPro; IPR006096; GLFV dehydrog C.  
 CC InterPro; IPR006097; GLFV dehydrog N.  
 CC Pfam; PF02008; GLFV dehydrog; 1.  
 CC Pfam; PF02812; GLFV dehydrog N; 1.  
 CC PRINTS; PR00082; GLFV dehydrog N.  
 CC PROSITE; PS00074; GLFV DEHYDROGENASE; 1.  
 CC Oxidoreductase; NAD; NADP; 3D-structure; Complete proteome.  
 KW ACT SITE 105 105  
 FT NP\_BIND 220 226 NAD (POTENTIAL).  
 FT NP\_BIND 220 226 AW -> WA (IN REF. 2).  
 FT CONFLICT 88 89 T -> K (IN REF. 2).  
 FT CONFLICT 366 366  
 FT TURN 5 5  
 FT HELIX 6 17  
 FT HELIX 18 20  
 FT HELIX 25 31  
 FT TURN 32 32  
 FT TURN 36 45  
 FT STRAND 47 48  
 FT TURN 51 62  
 FT STRAND 64 65  
 FT TURN 68 69  
 FT STRAND

```
FT STRAND 72 74
FT TURN 76 77
FT HELIX 80 96
FT TURN 97 98
FT STRAND 102 109
FT HELIX 112 114
FT HELIX 117 131
FT HELIX 132 134
FT STRAND 136 136
FT TURN 137 139
FT STRAND 140 141
FT STRAND 144 144
FT TURN 146 147
FT HELIX 150 164
FT TURN 165 166
FT HELIX 170 173
FT HELIX 179 181
FT TURN 182 182
FT TURN 185 189
FT HELIX 190 205
FT TURN 206 207
FT TURN 212 213
FT STRAND 215 219
FT HELIX 223 234
FT TURN 235 236
FT STRAND 239 244
FT STRAND 249 256
FT HELIX 258 268
FT TURN 272 273
FT TURN 275 276
FT STRAND 278 280
FT HELIX 282 287
FT STRAND 292 295
FT STRAND 302 302
FT TURN 304 305
FT HELIX 306 309
FT STRAND 314 316
FT STRAND 323 323
FT HELIX 325 333
FT TURN 334 335
FT STRAND 337 339
FT HELIX 341 344
FT TURN 345 346
FT HELIX 347 361
FT HELIX 367 391
FT TURN 392 393
FT HELIX 396 415
FT TURN 416 416
SQ SEQUENCE 420 AA; 47114 MW; 673DB20F8764A93C CRC64;

Query Match 7.1%; Score 87; DB 1; Length 420;
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 56; Conservative 29; Mismatches 73; Indels 58; Gaps 15;

QY 18 YLN-EALAWTFNDSNNYKTLGTSQITPALFPKAGDILYSLKDELGRTR-TARGTLTVAN 75
Db 147 YTNPQIMAWMDE-----YETISRRK-TPAFGIITGKPL-SIGGSLGRTEATARG----- 194

QY 76 VEGSYGVRQSGKQNPAGWTGNPNHVYKYLEWNLGLSYVG-----DFWNRSHLIA 126
Db 195 --ASYTIREA-----AKVLGW-----DTLKGKTIAGQYGNAGYVLAKIMSEDFGKMWAVS 244

QY 127 DSLGG-----DALRVNAVGTGTRQNVGGDQGGMRYTEQRAQEWLEARNRGYLYEVAPI 182
Db 245 DSKGGIYNPDGLNADEVILKWKREHSGSVKDFPGATNITN---EELLE----- 287

QY 183 YNADBLIPRAV--VVMOSSDNTINEKVLVYNTANG 216
Db 288 LEVDVLAAPAAIEEVIKKNADN-ICAKI-VAEVANG 321
```

```
ID DHE3_PYRHO STANDARD; PRT; 420 AA.
AC OS2310;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamate dehydrogenase (EC 1.4.1.3) (GDH).
GN GDHA OR GDH OR PH1593.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAI;
RA Gonzalez J.M., Robb F.T., Kato C.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Mauchi Y., Shizuwa H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76 (1998).
CC -1- CATALYTIC ACTIVITY: L-glutamate + H(2)O + NAD(P) (+) = 2-
oxoglutarate + NH(3) + NAD(P)H.
CC -1- SUBUNIT: Homohexamer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF035935; AAB99956.1; -.
DR HSP; P80319; IGTW.
DR InterPro; IPR006095; GLFV_dehydrog.
DR InterPro; IPR006096; GLFV_dehydrog_C.
DR InterPro; IPR006097; GLFV_dehydrog_N.
DR Pfam; PF00208; GLFV_dehydrog; 1.
DR Pfam; PF02812; GLFV_dehydrog_N; 1.
DR PRINTS; PR00082; GLFVDRGNASE.
DR PROSITE; PS00074; GLFV_DEHYDROGENASE; 1.
KW Oxidoreductase; NAD; NADP; Complete proteome.
FT ACT_SITE 105 105 BY SIMILARITY.
FT NP_BIND 220 226 NAD (POTENTIAL).
SQ SEQUENCE 420 AA; 47014 MW; 1198BEC2681B5AA2 CRC64;

Query Match 7.1%; Score 87; DB 1; Length 420;
Best Local Similarity 25.9%; Pred. No. 2.7;
Matches 56; Conservative 30; Mismatches 72; Indels 58; Gaps 15;

QY 18 YLN-EALAWTFNDSNNYKTLGTSQITPALFPKAGDILYSLKDELGRTR-TARGTLTVAN 75
Db 147 YTNPQIMAWMDE-----YETIARRK-TPAFGIITGKPL-SIGGSLGRNEATARG----- 194

QY 76 VEGSYGVRQSGKQNPAGWTGNPNHVYKYLEWNLGLSYVG-----DFWNRSHLIA 126
Db 195 --ASYTIREA-----AKVLGW-----LKGKTIAGQYGNAGYVLAKIMSEDFGKMWAVS 244

QY 127 DSLGG-----DALRVNAVGTGTRQNVGGDQGGMRYTEQRAQEWLEARNRGYLYEVAPI 182
Db 245 DSKGGIYNPDGLNADEVILKWKREHSGSVKDFPGA---TNISNELLE----- 287

RESULT 13
DHE3_PYRHO
```



```

Qy 183 YNADELIPRAV--VVSQSSDNTINEKVLVYNTANG 216
Db 288 LDVDVLAPAAIEEVITTKNADN-IRAKI-VAEVANG 321

RESULT 14
RPOB_HETCA
ID_RPOB_HETCA STANDARD; PRT; 1116 AA.
AC P36440;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6).
GN RPOB.
OS Heterosigma carterae.
OC Chloroplast.
OC Eukaryota; stramenopiles; Raphidophyceae; Heterosigma.
OX NCBI_TaxID=28465;
RN [1]
RP SEQUENCE FROM N.A.
RA Mangahas J.L., Cattolico R.A., Reynolds A.E.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: In chloroplasts, the RNA polymerase is composed of four
CC subunits: alpha, beta, beta', and beta".
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; X75815; CAA53450.1; -.
CC PIR; S41915; S41915.
CC HSSP; Q9KWU7; IHQM.
CC InterPro; IPR001572; RNA_pol_B.
CC Pfam; PF04563; RNA_pol_Rpb2_1; 1.
CC Pfam; PF04561; RNA_pol_Rpb2_2; 1.
CC Pfam; PF04565; RNA_pol_Rpb2_3; 1.
CC Pfam; PF00562; RNA_pol_Rpb2_6; 1.
CC Pfam; PF04560; RNA_pol_Rpb2_7; 1.
CC PROSITE; PS01166; RNA_POL_BETA; 1.
CC Transference; DNA-directed RNA polymerase; Transcription; Chloroplast.
CC SEQUENCE 1116 AA; 125818 MW; EC6C83C81234435B CRC64;

Query Match 7.1%; Score 86.5; DB 1; Length 1116;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 55; Conservative 19; Mismatches 79; Indels 75; Gaps 12;

Qy 6 SNDVLND-GASKYLNALAWTFNDSPNYKTLGTSGITPALPKAGDILYSK-----LD 59
Db 831 SVDILNPLGVPSRN-----VGQIFECLLGLAGDQLNKRKILPFD 872

Qy 60 ELGRTRTARGTLTYANVEGSGVRSF---GKNQNP-----GWTGNPNH 101
Db 873 EM-----YQNEASRLINQKDAKQKQNPWLSAYSPGKILLSDGRTGE--- 918

Qy 102 VKYKIEWLNGSLVGVDFWNRSHLIADSLGGDALARNAVTTGTQNVGGRDQKGMRYTEQ 161
Db 919 -KFDNPLVGRSYI---LKLHLVEDKI--HARSTGYSLLITQQPVGKGQNGGQRFGE 972

Qy 162 RAQEM-LEANNPDGVLVYEVAPIYNADELI PRAVVVVSQSSDNTINEKV 208
Db 973 --EVMLEAFGAAYTLOELLTIKSD-----MQGRDDVLNIV 1008

RESULT 15
CHDI_HUMAN
ID_CHDI_HUMAN STANDARD; PRT; 1709 AA.
AC O14636;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chromodomain-helicase-DNA-binding protein 1 (CHD-1).
GN CHD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97470991; PubMed=9326634;
RA Woodage T., Baerai M.A., Baxevanis A.D., Hieter P., Collins F.S.;
RT "Characterization of the CHD family of proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11472-11477(1997).
CC -1- FUNCTION: SEQUENCE-SELECTIVE DNA-BINDING PROTEIN. COULD PLAY AN
CC IMPORTANT ROLE IN GENE REGULATION.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
CC -1- SIMILARITY: Contains 2 chromo domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AF006513; AAB87381.1; -.
CC HSSP; P23197; IAP0.
CC Genew; HGNC:1915; CHD1.
CC MIM; 602118; -.
CC GO; GO:0004003; F:ATP dependent DNA helicase activity; TAS.
CC GO; GO:0007001; P:chromosome organization and biogenesis (gen. .; TAS.
CC GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
CC InterPro; IPR000953; Chromo.
CC InterPro; IPR001410; DEAD.
CC InterPro; IPR001650; Helicase_C.
CC InterPro; IPR000330; SNF2_N.
CC Pfam; PF00385; chromo; 2.
CC Pfam; PF00271; helicase_C; 1.
CC Pfam; PF00176; SNF2_N; 1.
CC SMART; SM00298; CHROMO; 2.
CC SMART; SM00487; DEXDC; 1.
CC SMART; SM00490; HELIC; 1.
CC PROSITE; PS00598; CHROMO_1; 2.
CC PROSITE; PS00013; CHROMO_2; 2.
CC DNA-binding; ATP-binding; Helicase; Nuclear protein; Repeat.
KW DOMAIN 1 70 SER-RICH.
FT DOMAIN 117 137
FT DOMAIN 272 364 CHROMO 1.
FT DOMAIN 389 452 CHROMO 2.
FT NP_BIND 506 513 ATP (POTENTIAL).
FT SITE 614 617 DEAH BOX.
FT DOMAIN 1628 1644 3 X 5 AA REPEATS OF H-S-D-H-R.
FT REPEAT 1628 1632 1.
FT REPEAT 1634 1638 2.
FT REPEAT 1640 1644 3.
SQ SEQUENCE 1709 AA; 196517 MW; 416409C913D6A935 CRC64;

Query Match 7.1%; Score 86.5; DB 1; Length 1709;
Best Local Similarity 20.5%; Pred. No. 18;
Matches 30; Conservative 23; Mismatches 62; Indels 31; Gaps 5;

Qy 61 LGTRTRTARGTLTYANVEGSGVRSQSGKQNPAGTGNPHVVKYKIEWLNGSLVGVDFWN 120
Db 282 IGRKGATGATTTIYAVEADGDPNAGFEKKN-----EPGEIQYLKWK-KGWSHIHTWE 333

```

Qy	121	RSHLIADSLGGDALRVNAVTTGTNTONVGRDQKGMRYTEORAEWLE-ANRDGYLYYE-	178
Db	334	TEETL-----KQNVNRMKKLDNYKKDKQETKRWLKNASPDVEYINC	376

Db 334 TEETL-----KQQNVRGMKKLDNYKKKDQETKRWLKNASPEDVEYYNC 376

QY 179 ----VAPIYNADELIPRAVVSMQSS 200

DB 377 QQELTDDLHKQYQIVGRIIAHSNQKS 402

Search completed: January 5, 2004, 18:39:13  
Job time : 20.3996 secs

Job time : 20.3996 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 16:40:38 ; Search time 53.1381 Seconds  
(without alignments)  
1112.084 Million cell updates/sec

Title: US-08-482-785-9  
Perfect score: 1220  
Sequence: 1 RQGVNDVNDGASKYLN.....VYNTANGTYINHGPTQK 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23.3\*  
1: sp\_archaea\*  
2: sp\_bacteria\*  
3: sp\_fungi\*  
4: sp\_human\*  
5: sp\_invertebrate\*  
6: sp\_mammal\*  
7: sp\_mhc\*  
8: sp\_organelle\*  
9: sp\_phage\*  
10: sp\_plant\*  
11: sp\_rodent\*  
12: sp\_virus\*  
13: sp\_vertebrate\*  
14: sp\_unclassified\*  
15: sp\_rvirus\*  
16: sp\_bacteriap\*  
17: sp\_archaeap\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1220	100.0	271	16 Q9ACN6	Q9ACN6 streptococc
2	1216	99.7	271	16 Q54969	Q54969 streptococc
3	719.5	59.0	264	16 Q8K6X0	Q8K6X0 streptococc
4	401	32.9	332	16 Q8NZT6	Q8NZT6 streptococc
5	391.5	32.1	385	2 Q33735	Q33735 streptococc
6	372	30.5	327	2 Q93CC1	Q93CC1 streptococc
7	363	29.8	326	16 Q8K6B9	Q8K6B9 streptococc
8	181.5	14.9	278	2 Q93NM6	Q93NM6 lactobacill
9	179.5	14.7	252	2 Q8VQ29	Q8VQ29 streptococc
10	173.5	14.2	266	16 Q8CM17	Q8CM17 streptococc
11	171.5	14.1	252	16 Q9A0M1	Q9A0M1 streptococc
12	158.5	13.0	266	16 Q8P0B6	Q8P0B6 streptococc
13	158.5	13.0	268	16 Q99Z26	Q99Z26 streptococc
14	133.5	12.6	263	16 Q9CJ50	Q9CJ50 lactococcu
15	150	12.3	343	16 Q8E7K1	Q8E7K1 streptococc
16	134	11.0	354	16 Q8E6F9	Q8E6F9 streptococc

17	134	11.0	354	16 Q8E0U1	Q8E0U1 streptococc
18	131	10.7	261	16 Q8E6B6	Q8E6B6 streptococc
19	131	10.7	261	16 Q8E0P2	Q8E0P2 streptococc
20	128	10.5	285	16 Q8E5U4	Q8E5U4 streptococc
21	127	10.4	285	16 Q8E0E7	Q8E0E7 streptococc
22	118.5	9.7	287	2 Q8VT23	Q8VT23 enterococcu
23	112.5	9.2	293	16 Q8DT59	Q8DT59 streptococc
24	106	8.7	1233	16 Q8NSE4	Q8NSE4 corynebacte
25	105	8.6	24	2 Q9R4B2	Q9R4B2 streptococc
26	104.5	8.6	455	16 Q9Z1E1	Q9Z1E1 helicobacte
27	104.5	8.6	764	16 Q8XTM0	Q8XTM0 ralstonia s
28	104	8.5	21	2 Q9R5K0	Q9R5K0 streptococc
29	103.5	8.5	705	16 Q8XTL2	Q8XTL2 ralstonia s
30	102.5	8.4	353	10 Q9LR37	Q9LR37 arabidopsis
31	100	8.2	160	16 Q9A0I2	Q9A0I2 streptococc
32	94.5	7.7	468	16 Q8XSL5	Q8XSL5 ralstonia s
33	94	7.7	319	16 Q8NX49	Q8NX49 staphylococ
34	93.5	7.7	74	2 Q8GCX6	Q8GCX6 enterococcu
35	93.5	7.7	561	16 Q97TK3	Q97TK3 clostridium
36	92.5	7.6	372	16 Q8G2H7	Q8G2H7 brucella su
37	92.5	7.6	763	16 Q8P5Q7	Q8P5Q7 xanthomonas
38	92	7.5	293	2 Q9ZAK0	Q9ZAK0 clostridium
39	92	7.5	867	2 Q93G68	Q93G68 escherichia
40	91.5	7.5	788	16 Q8XGB7	Q8XGB7 ralstonia s
41	91.5	7.5	1118	2 Q9S0X0	Q9S0X0 clostridium
42	91.5	7.5	1118	2 Q9X721	Q9X721 clostridium
43	91.5	7.5	1534	5 Q26946	Q26946 trypanosoma
44	91	7.5	386	16 Q92ES3	Q92ES3 listeria in
45	91	7.5	1429	2 Q9LAP7	Q9LAP7 alteromonas

## ALIGNMENTS

### RESULT 1

ID	Q9ACN6	PRELIMINARY;	PRT;	271 AA.
AC	Q9ACN6;			
DT	01-JUN-2001 (Tremblrel. 17, Created)			
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT	01-OCT-2002 (Tremblrel. 22, Last annotation update)			
DE	Streptodornase B precursor (Mitogenic factor 25K).			
GN	DNASEB OR MF OR SPYM3_1745.			
OS	Streptococcus pyogenes, and			
OS	Streptococcus pyogenes (serotype M3).			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1314, 198466;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=S.pyogenes; STRAIN=C203S;			
RX	MEDLINE=21233096; PubMed=11335140;			
RA	Garlach D., Schmidt K.H., Fleischer B.;			
RT	"Basic streptococcal superantigens (SPEX/SMEZ or SPEC) are responsible			
RL	for the mitogenic activity of the so called mitogenic factor MF.";			
RL	FEMS Immunol. Med. Microbiol. 30:209-216(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=S.pyogenes; STRAIN=MCAS315 / Serotype M3;			
RX	MEDLINE=22133808; PubMed=12122206;			
RA	Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,			
RA	Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,			
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,			
RA	Schlievert P.M., Musser J.M.;			
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus:			
RT	phage-encoded toxins, the high-virulence phenotype, and clone			
RT	emergence.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).			
RL	EMBL; AJ295272; CAC35734.1; -.			
DR	EMBL; AF014170; AM80352.1; -.			
KW	Signal; Complete proteome.			
FT	SIGNAL 1 42			POTENTIAL.
FT	CHAIN 43 271			STREPTODORNASE B.

```
SQ SEQUENCE 271 AA; 30090 MW; 8FD849E3CD01352 CRC64;
Query Match 100.0%; Score 1220; DB 16; Length 271;
Best Local Similarity 100.0%; Pred. No. 4e-92; Indels 0; Gaps 0;
Matches 229; Conservative 0; Mismatches 0;
Qy 1 RQTQVSDVNDGASKYLNEALAWTFDPSNPKYKLTGTSQITPALPKAGDILYSLKDE 60
Db 43 RQTQVSDVNDGASKYLNEALAWTFDPSNPKYKLTGTSQITPALPKAGDILYSLKDE 102
Qy 61 LGRTTRTARGTLTYANVEGSGVGRQSGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 120
Db 103 LGRTTRTARGTLTYANVEGSGVGRQSGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 162
Qy 121 RSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 180
Db 163 RSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 222
Qy 181 PIYNADELIPRAVVSQSSDNTINEKVLVNTANGTYTINHNGTPTQK 229
Db 223 PIYNADELIPRAVVSQSSDNTINEKVLVNTANGTYTINHNGTPTQK 271
RESULT 2
Q54969
ID Q54969 PRELIMINARY; PRT; 271 AA.
AC Q54969
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Mitogenic factor precursor (Mitogenic factor, 25K).
GN MF OR SPY2043 OR SPYM18 2104.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NY-5;
RX MEDLINE=94009636; PubMed=8405402;
RA Iwasaki M., Igarashi H., Hinuma Y., Yutsudo T.;
RT "Cloning, characterization and overexpression of a Streptococcus
RT pyogenes gene encoding a new type of mitogenic factor.";
RL FEBS Lett. 331:187-192 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SV;
RX Hong K.;
RA "A novel cloning method used arbitrarily primed PCR.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
```

```
DR EMBL; D13428; BAA02693.1; -.
DR EMBL; AB030578; BAB16025.1; -.
DR EMBL; AE006625; AAK34710.1; -.
DR EMBL; AE010113; AAL98563.1; -.
KW Nuclease; Signal; Complete proteome.
FT SIGNAL 1 43 POTENTIAL.
FT CHAIN 44 271 MITOGENIC FACTOR.
SQ SEQUENCE 271 AA; 30062 MW; 8FD89DA38A124352 CRC64;
Query Match 99.7%; Score 1216; DB 16; Length 271;
Best Local Similarity 99.6%; Pred. No. 8.6e-92;
Matches 228; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 RQTQVSDVNDGASKYLNEALAWTFDPSNPKYKLTGTSQITPALPKAGDILYSLKDE 60
Db 43 RQTQVSDVNDGASKYLNEALAWTFDPSNPKYKLTGTSQITPALPKAGDILYSLKDE 102
Qy 61 LGRTTRTARGTLTYANVEGSGVGRQSGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 120
Db 103 LGRTTRTARGTLTYANVEGSGVGRQSGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 162
Qy 121 RSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 180
Db 163 RSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 222
Qy 181 PIYNADELIPRAVVSQSSDNTINEKVLVNTANGTYTINHNGTPTQK 229
Db 223 PIYNADELIPRAVVSQSSDNTINEKVLVNTANGTYTINHNGTPTQK 271
RESULT 3
Q8K6X0
ID Q8K6X0 PRELIMINARY; PRT; 264 AA.
AC Q8K6X0
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Putative mitogenic factor-phase associated.
GN MF4 OR SPYM3 1095.
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phase-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL; AE014158; AAM79702.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 264 AA; 29333 MW; 6E3954B68A309399 CRC64;
Query Match 59.0%; Score 719.5; DB 16; Length 264;
Best Local Similarity 62.1%; Pred. No. 4.2e-51;
Matches 141; Conservative 25; Mismatches 56; Indels 5; Gaps 2;
Qy 1 RQTQVSDVNDGASKYLNEALAWTFDPSNPKYKLTGTSQITPALPKAGDILYSLKDE 60
Db 36 KEVNVANRY---NSSDTYLPALSLTLETSPNPKYKLVGESIVENLPPPKGQIVYGGDS 92
Qy 61 LGRTTRTARGTLTYANVEGSGVGRQSF--GKNQNPAGTGNPNHVKYKIEWLNGLSYVGDF 118
Db 93 LGRTLTRVGRITLVNVLGSGVNIKDFKRSKAEITSLGWLGNKNGEVYVYKGLGDDSYQGYF 152
Qy 119 WNRSHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYE 178
```

```

Db 153 WNKSHLIADSLGGDALRVNAITGTRTQNVGSGNGGWRYTEIKSQKLEAHRDGYLYE 212
Qy 179 VAPYNADLIPRAVVMQSSDNTINEKVLVYNTANGYTYNYHNGT 225
Db 213 AMPIYQGNELVPRVVSVLSSDNTINEKVIYVNVANGYTYIDYNOGT 259

RESULT 4
Q8NZT6
ID Q8NZT6 PRELIMINARY; PRT; 332 AA.
AC Q8NZT6;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Streptococcaceae.
GN SPYM18.1746.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylvia G.L., Sturdevant D.B., Rickielfs S.M., Porcella S.F.,
RA Parkins L.D., Bares S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Mueser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AB010084; AAL98274.1; -.
KW Complete proteome.
SQ SEQUENCE 332 AA; 37652 MW; 58DC9F930865FBC0 CRC64;

Query Match 32.9%; Score 401; DB 16; Length 332;
Best Local Similarity 36.9%; Pred. No. 7.5e-25;
Matches 90; Conservative 32; Mismatches 78; Indels 44; Gaps 4;

Qy 25 WFNDSPNYKTLGTSQITPALFP-----KAG-----DILYSKLDLGR 64
Db 46 WNVQHNYIVVEGSHLNKEKFPQIYHTEKVKRSGQSTKRVTSIDYIGVGRS 105
Qy 65 RTAGTLTYANVEGSGVRSFGKNQNPAGWTG-----NP-----N 100
Db 106 GEAVGVITKMDMSAGYREKWKESKPEPSGMYFFKNTNQRTATESDYKHSKPNVSKISN 165
Qy 101 HVKYKIEWLNGLSVGVDFWNRSHLIADSLGGDALRVNAITGTRTQNVGGRDQGGWRYTE 160
Db 166 NIKASIPLSNGRTRHGFLDFRSHLIADSLGGRPRNNLITGTRTQNVGNDRKGMQYIE 225
Qy 161 QRAQEWLEARNRGGVLYVEVAPIYNADLIPRAVVMQSSDNTINEKVLVYNTANGYTYN 220
Db 226 NKVLDHIKQPKVHVYKATPVYQGSLLPRVLVLSALSDGTDITETRVFVNAVAGFNID 285
Qy 221 YHNG 224
Db 286 YQNG 289

RESULT 5
Q33735
ID Q33735 PRELIMINARY; PRT; 385 AA.
AC Q33735;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE Streptococcaceae.
GN SDA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```

```

OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RA Podbielski A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=serotype M49;
RX MEDLINE=97101060; PubMed=8945587;
RA Podbielski A., Zarges I., Flosdorff A., Weber-Heynemann J.;
RT "Molecular characterization of a major serotype M49 group A
RT streptococcal DNase gene (sdnD).";
RL Infect. Immun. 64:5349-5356(1996).
DR EMBL; X84793; CAA59264.1; -.
SQ SEQUENCE 385 AA; 43476 MW; 14922B0AFC3BD6D5 CRC64;

Query Match 32.1%; Score 391.5; DB 2; Length 385;
Best Local Similarity 35.0%; Pred. No. 5.5e-24;
Matches 91; Conservative 38; Mismatches 84; Indels 47; Gaps 5;

Qy 12 NDGASKYLNEALA---WTFNDSPNYKTLGTSQI-----TPALFPKAG----- 51
Db 30 NTSSTIYAESATISNNWSIEQHFNHYHVEGKAQLDIKNFPPELYRTTERTVYKSGSQSTKPV 89
Qy 52 ---DILYSKLDLGRTRTARGTLTYANVEGSGVRSFGKNQNPAGWTG----- 97
Db 90 TVSNHYSLVDGYRSGEAYGIITKMDMSAGYREKWKESKPEPSGMYFFKNTNQRT 149
Qy 98 -----NP-----NHVKYKIEWLNGLSVGVDFWNRSHLIADSLGGDALRVNAITGTRT 144
Db 150 ESDYKHSKPNVSKISNNIKASILLNSGNVRNGVLFDRSHLIADSLGGRPRNNLITGTRT 209
Qy 145 QNVGGRDQGGWRYTEQRAQEWLEARNRGGVLYVEVAPIYNADLIPRAVVMQSSDNTI 204
Db 210 QNVGNDRKGMQYIENKVLDHIKRNPKVHVYKATPVYQGSLLPRVLVLSALSDGFI 269
Qy 205 NEKVLVYNTANGYTYNYHNG 224
Db 270 DETRVFVNAVAGFNIDYQNG 289

RESULT 6
Q93CC1
ID Q93CC1 PRELIMINARY; PRT; 327 AA.
AC Q93CC1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Sdalpha deoxyribonuclease.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Hasegawa T., Torii K., Hashikawa S., Iinuma Y., Ohta M.;
RT "Cloning and characterization of the deoxyribonuclease sdalphi gene
RT from Streptococcus pyogenes.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410852; AAL02116.1; -.
SQ SEQUENCE 327 AA; 36834 MW; 62DDD3443017071F CRC64;

Query Match 30.5%; Score 372; DB 2; Length 327;
Best Local Similarity 35.0%; Pred. No. 1.7e-22;
Matches 85; Conservative 34; Mismatches 78; Indels 46; Gaps 4;

Qy 26 TFDNDSPNYKTLGTSQITPALFPK-----AGDILYSKLDLGRTR 65
Db 43 TYGEYKDYVYIGESNIDQSAFPKTYKTYRVYKGGTSEKVTVDVYVNPDLGDKRST 102
Qy 66 TARGTLTYANVEGSGVRSFGKNQNPAGW-----TGN-----PNH 101

```

```
Db 103 GAYGVVTKDMIDMSKGYREKWEINTEPNPESGWFRFYNRADNEEISEKEYDSRRTSKYKVTNN 162
Qy 102 VKYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNAVTCRTONVGGDQKGMRYTEQ 161
Db 163 VPVVLTLKQYNSHLFVSHLFPADSLGSKSRKNAITGTQONVGT--KGMQYIEK 220
Qy 162 RAQEWLEARNRGYLYEYVAPYNADELIPRAVVVSMQSDNTINEKVLVNTANGTYTNY 221
Db 221 KVLSHITKNDVYVYFVSAIPYQGTETLLARSVLVSALSSDGVINEVTVFNTADGNINY 280
Qy 222 HNG 224
Db 281 EKG 283

RESULT 7
Q8K6B9 ID Q8K6B9 PRELIMINARY; PRT; 326 AA.
AC Q8K6B9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
GN Sdn OR SPYM3 1409
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=198466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;
RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence."
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).
DR EMBL; AF014163; AAM80016.1; -.
KW Complete proteome.
SQ SEQUENCE 326 AA; 36645 MW; C9635F8A0CA6A51F CRC64;

Query Match 29.8%; Score 363; DB 16; Length 326;
Best Local Similarity 34.6%; Pred. No. 9.5e-22;
Matches 84; Conservative 34; Mismatches 79; Indels 46; Gaps 4;

Qy 26 TENDSPNYKTLGTQITPALPPK-----AGDILYKLDLGRTR 65
Db 43 TYGEYKDYTVIGESNIDQSAPPKLYKTERVYKGGTSEKRVTVSDVYNPLDGYKRST 102
Qy 66 TARGTLTYANVEGSGVRSFGKQNPAGW-----TGN-----PNH 101
Db 103 GAYGVVTKDMIDMSKGYREKWEINTEPNPESGWFRFYNRADNEEISEKEYDSRRTSKYKVTNN 162
Qy 102 VKYKIEWLNGLSYVGDFFWNRSHLIADSLGGDALRVNAVTCRTONVGGDQKGMRYTEQ 161
Db 163 VPVVLTLKQYNSHLFVSHLFPADSLGSKSRKNAITGTQONVGT--KGMQYIEK 220
Qy 162 RAQEWLEARNRGYLYEYVAPYNADELIPRAVVVSMQSDNTINEKVLVNTANGTYTNY 221
Db 221 KVLSHITKNDVYVYFVSAIPYQGTETLLARSVLVSALSSDGVINEVTVFNTADGNINY 280
Qy 222 HNG 224
Db 281 EKG 283

RESULT 8
Q93MN6 ID Q93MN6 PRELIMINARY; PRT; 278 AA.
AC Q93MN6;
```

```
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
GN Hypothetical membrane protein LsaO.
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OX NCBI_TaxID=1599;
RN [1]
RP SEQUENCE FROM N.A.
RA Duzet A.-M., Chaillou S., Hissler L., Stentz R., Champomier-Verges M.,
RA Albert C.-A., Zagorec M.;
RT "Physical and genetic map of the Lactobacillus sakei 23K chromosome."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF400065; AAK92009.1; -.
SQ SEQUENCE 278 AA; 30597 MW; DBA1337649DEB8AD CRC64;

Query Match 14.9%; Score 181.5; DB 2; Length 278;
Best Local Similarity 27.1%; Pred. No. 5.8e-07;
Matches 62; Conservative 33; Mismatches 77; Indels 57; Gaps 10;

Qy 4 QVSNVDVINDGASKYLNEALAWTFNDSPNYKTLGTSTQITPALFPKAGDILYKLDLGR 63
Db 95 QGDNEIIVNNQPTFTTADLATT--NGP--WQT-----FSNLDQLNR 132
Qy 64 TRTARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVYKIEWLNGLSYVGDFFWNRSH 123
Db 133 AGTANALLN---KSMPTAKREGLTWNPTGW---RNKRVKSGWL-----YNRSH 175
Qy 124 LIADSLGGDALR-VNAVTCRTONVGGDQKGMRYTEQRAQ-----WLEARNRGYLYY 177
Db 176 LIGYQITGENNPNKLUITGRQLNA-----PEMLAHESDIAAYLKXHLSDVRY 224
Qy 178 EVAPIYNADELIPRAVVVSMQSDNTINEKVLVNTANGTYTNYHNGT 225
Db 225 RVTPIFRGNELLARGVQMEGSGSDNSIQFNFTYIFNVADGVTLNADGS 273

RESULT 9
Q8VQ29 ID Q8VQ29 PRELIMINARY; PRT; 252 AA.
AC Q8VQ29;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE M2 variant.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RA Hasegawa T., Torii K., Hashikawa S., Iinuma Y., Ohta M.;
RT "Cloning and characterization of two new novel DNases from
RT Streptococcus pyogenes."
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF461500; AAL66371.1; -.
SQ SEQUENCE 252 AA; 28347 MW; 92CABEC832F5CFB8 CRC64;

Query Match 14.7%; Score 179.5; DB 2; Length 252;
Best Local Similarity 28.0%; Pred. No. 7.4e-07;
Matches 63; Conservative 37; Mismatches 80; Indels 45; Gaps 13;

Qy 33 YKTLGTSTQITPALFPKAGDILYKLDLGRTRTARGTLTYANVEGSGVRSFGK-NQN 91
Db 44 HYKNTASSKLIP--FTANYQLQLGELDNLN-----RATFSHIQLQDRHETKDVTKINYD 96
Qy 92 PAGWTGNPNHVYKIEWLNG--LSVVGDFWNRSHLIADSLGG--DALRVNAVTCRTONV 147
Db 97 PVGW-----H-NYQFPYGDGSKSWV---MNRGHLVGQFCGLNDEPR-NLVTMTAMLT 146
Qy 148 G---GRDQKG--GMRYTEQRAQEWLEARNRGYLYEYVAPYNADELIPRAV----- 193
```

```
Db 147 GAYSGANDSPEGMLYENRLDSWLAHPDFWLDYKVTPIYSGNEVVPRQIELQYVGIDS 206
Qy 194 -----VVSQSSDNTINEK-----VLVYNTANGTYTINYHNGTPTQK 229
Db 207 SGBLLTIRLNSKESIDENGVTTVILENSAPNILDYLANGTATPK 251

RESULT 10
Q8CM17
ID Q8CM17 PRELIMINARY; PRT; 266 AA.
AC Q8CM17
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN GBS0382 OR GBS0712 OR GBS0997.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Zouinok C., Buchrieser C., Chevalier P., Frangeul L.,
RA Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
RL Mol. Microbiol. 45:1499-1513 (2002).
DR EMBL; AL766845; CAD46026.1; -.
DR EMBL; AL766846; CAD46356.1; -.
DR EMBL; AL766848; CAD46656.1; -.
DR Sagaliet; gbs0382; -.
DR Sagaliet; gbs0712; -.
DR Sagaliet; gbs0997; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 266 AA; 30394 MW; D07DEEF65BF006A1 CRC64;

Query Match 14.2%; Score 173.5; DB 16; Length 266;
Best Local Similarity 31.2%; Pred. No. 2.5e-06;
Matches 54; Conservative 21; Mismatches 83; Indels 15; Gaps 7;

Qy 55 YSKLDLGRTRTARGTLTYANVSGYGVQSGFNQNPAGWGNPNHVKYKIEWLNGSLY 114
Db 99 YDNLDFLRVGVAEMLGKELMPKE--ARQDI-SSVKPTGKNK-----KITFNGKQDY 149
Qy 115 VGFWRSHLIADSLGDALRV-NAVGTGRTQNVGRDQKGMRYTEQRAQEWLEARNRG 173
Db 150 L---YNRSHLIGFQSGENANVKNLFTGTRALNANFNDDKSSMVYVENEVANYIK-KTNH 205
Qy 174 YLYVEAPIYNADELIPRAVVVSQSS-SDNTINEKVLVNTANGTYTINYHNGT 225
Db 206 HVRYRTPLPKVELVARGVRIEAQSIETDISFDVIYFNGQPGYDIDYLTGS 258

RESULT 11
Q9A0M1
ID Q9A0M1 PRELIMINARY; PRT; 252 AA.
AC Q9A0M1
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Putative Dnaase (Similar to mitogenic factor), phage associated
DE (MF2).
GN MF2 OR SPY0712 OR SPYM18_0779.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Sylva G.L., Burdian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AB006523; AAK33665.1; -.
DR EMBL; AB010008; AAL97446.1; -.
KW Complete proteome.
SQ SEQUENCE 252 AA; 28373 MW; 71EA2332E358440D CRC64;

Query Match 14.1%; Score 171.5; DB 16; Length 252;
Best Local Similarity 27.6%; Pred. No. 3.4e-06;
Matches 62; Conservative 37; Mismatches 81; Indels 45; Gaps 13;

Qy 33 YKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVSGYGVQSGFK-NQN 91
Db 44 HYKNTVSSKLLP--FTANYQLQLGELDNLN-----RATFSHIQLQDRHETKQVTRKINYD 96
Qy 92 PAGWTGNPNHVKYKIEWLNG--LSVYVGDFFWRSHLIADSLG--DALRVNAVGTGRTQNV 147
Db 97 PVGW-----H-NYQFPYGDGKSSWV---MNRGLHVGVOFCGLNDEPR-NLVAMTAWLNT 146
Qy 148 G---GRDQKG--GMRYTEQRAQEWLEARNRGDGLYVEAPIYNADELIPRAV----- 193
Db 147 GAYSGANDSPEGMLYENRLDSWLAHPDFWLDYKVTPIYSGNEVVPRQIELQYVGIDS 206
Qy 194 -----VVSQSSDNTINEK-----VLVYNTANGTYTINYHNGTPTQK 229
Db 207 SGBLLTIRLNSKESIDENGVTTVILENSAPNILDYLANGTATPK 251

RESULT 12
Q8P0B6
ID Q8P0B6 PRELIMINARY; PRT; 266 AA.
AC Q8P0B6
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE MF3.
GN SPYM18_1446.
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Sylva G.L., Burdian K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
DR EMBL; AB010063; AAL98024.1; -.

```

```
KW Complete proteome.
SQ SEQUENCE 266 AA; 30153 MW; 45475BC64C408F9A CRC64;

Query Match
Best Local Similarity 13.0%; Score 158.5; DB 16; Length 266;
Matches 55; Conservative 26; Mismatches 65; Indels 31; Gaps 12;

QY 37 LGTSQI-TPALFPKAG--DILYSKLDLGRTRTARGTLTYANV-----EGSYGVRSFGKN 89
DB 45 LGTRATETPGILPTGSGVQLVGLDNLQRP-----TPAHQLKQDDEPNIKRK-GLK 96
QY 90 QNPAGWTGNPNHVYKIEWNLGLSYVGDFW--NRSHLIADSLGG--DALRVNAVVTGRTQ 145
DB 97 FNPFGW-----H-NYKLTDAWGKT---TWLMDRGHLVGVQFSLNDEPK-NLVTMTKYL 145
QY 146 NVGGRDQKG-CMRYTEQRAQEWLEARNRQDGLYVEVAPIYNADLIPRAVVMQSSD 201
DB 146 NTGFSQDNPLGMLYENRLDSWLALHPNFWLDYKVTVPVHKNELVPRQVVLQVVGID 202

RESULT 13
Q99Z26 PRELIMINARY; PRT; 268 AA.
ID AC Q99Z26
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
DR Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
KW EMBL; AE006579; AAK34241.1; -.
SQ SEQUENCE 268 AA; 30398 MW; 24EA4762A5DDAA5E CRC64;

Query Match
Best Local Similarity 13.0%; Score 158.5; DB 16; Length 268;
Matches 55; Conservative 26; Mismatches 65; Indels 31; Gaps 12;

QY 37 LGTSQI-TPALFPKAG--DILYSKLDLGRTRTARGTLTYANV-----EGSYGVRSFGKN 89
DB 47 LGTRATETPGILPTGSGVQLVGLDNLQRP-----TPAHQLKQDDEPNIKRK-GLK 98
QY 90 QNPAGWTGNPNHVYKIEWNLGLSYVGDFW--NRSHLIADSLGG--DALRVNAVVTGRTQ 145
DB 99 FNPFGW-----H-NYKLTDAWGKT---TWLMDRGHLVGVQFSLNDEPK-NLVTMTKYL 147
QY 146 NVGGRDQKG-CMRYTEQRAQEWLEARNRQDGLYVEVAPIYNADLIPRAVVMQSSD 201
DB 148 NTGFSQDNPLGMLYENRLDSWLALHPNFWLDYKVTVPVHKNELVPRQVVLQVVGID 204

RESULT 14
Q9CJ50 PRELIMINARY; PRT; 263 AA.
ID AC Q9CJ50
RC STRAIN=21192684; PubMed=11296296;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
DR Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
KW EMBL; AE006579; AAK34241.1; -.
SQ SEQUENCE 263 AA; 28854 MW; EAA95B6FA317207C CRC64;

Query Match
Best Local Similarity 12.6%; Score 153.5; DB 16; Length 263;
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;

QY 26 TFNDSPNYKTLGTS-QITPALFPKAGDILYSKLDLGR-----RTARGTL 71
DB 39 TDSSQSVSTKSLASSVKQAPLTFKNQRMVMAWNTDALGRAVDSHIQKDSQEPKVRPEL 98
QY 72 TYANVEGSGVRSFGKNQNPAGWTGNPNHVYKIEWNLGLSYVGDFW--NRSHLIA--- 126
DB 99 TY-----H-NYNYFYKKSQSGSGIKRWMLMARGHLVGVQF 135
QY 127 DSLGGDALRV-----NAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLY 177
DB 136 SGLNNEARNLVPETAMFNGNGTGTNDGNT-----ASMLYENRLDSWLALHPN 189
QY 178 EVAPIYNADLIPRAV---VSMQSSDNTINPK-----VLTNTANGYTI 219
DB 190 QVTPLYEGNELLPRQIRLAYVIGIDKNGQTLISIKLGGREKSGNGCATVVLVDNAPNAKI 249
QY 220 NYHNGT 225
DB 250 NYADGT 255

RESULT 15
Q8E7K1 PRELIMINARY; PRT; 343 AA.
ID AC Q8E7K1
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,
RA Msadek T., Zouine M., Couve E., Lallouli L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae, a pathogen causing
RT invasive neonatal disease.";
DR Mol. Microbiol. 45:1499-1513(2002).
KW EMBL; AL766844; CAD45798.1; -.
DR Sagalish; gbs0153; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 343 AA; 38877 MW; 11BF35A12D959DF3 CRC64;

Query Match
Best Local Similarity 12.3%; Score 150; DB 16; Length 343;
Matches 62; Conservative 33; Mismatches 102; Indels 46; Gaps 11;
```

```
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
ON NCBI_TaxID=1360;
RX STRAIN=IL1403;
RC SEQUENCE FROM N.A.
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006253; AAK04254.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 263 AA; 28854 MW; EAA95B6FA317207C CRC64;
```

```
Query Match
Best Local Similarity 12.6%; Score 153.5; DB 16; Length 263;
Matches 60; Conservative 26; Mismatches 85; Indels 75; Gaps 11;

QY 26 TFNDSPNYKTLGTS-QITPALFPKAGDILYSKLDLGR-----RTARGTL 71
DB 39 TDSSQSVSTKSLASSVKQAPLTFKNQRMVMAWNTDALGRAVDSHIQKDSQEPKVRPEL 98
QY 72 TYANVEGSGVRSFGKNQNPAGWTGNPNHVYKIEWNLGLSYVGDFW--NRSHLIA--- 126
DB 99 TY-----H-NYNYFYKKSQSGSGIKRWMLMARGHLVGVQF 135
QY 127 DSLGGDALRV-----NAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRQDGLY 177
DB 136 SGLNNEARNLVPETAMFNGNGTGTNDGNT-----ASMLYENRLDSWLALHPN 189
QY 178 EVAPIYNADLIPRAV---VSMQSSDNTINPK-----VLTNTANGYTI 219
DB 190 QVTPLYEGNELLPRQIRLAYVIGIDKNGQTLISIKLGGREKSGNGCATVVLVDNAPNAKI 249
QY 220 NYHNGT 225
DB 250 NYADGT 255
```

```
RESULT 15
Q8E7K1 PRELIMINARY; PRT; 343 AA.
ID AC Q8E7K1
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevalier F., Prangeul L.,
RA Msadek T., Zouine M., Couve E., Lallouli L., Poyart C., Trieu-Cuot P.,
RA Kunst F.;
RT "Genome sequence of Streptococcus agalactiae (serotype III).
RT Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
ON NCBI_TaxID=216495;
RX STRAIN=IL1403;
RC SEQUENCE FROM N.A.
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
DR EMBL; AE006253; AAK04254.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 263 AA; 28854 MW; EAA95B6FA317207C CRC64;
```

```
Query Match
Best Local Similarity 12.3%; Score 150; DB 16; Length 343;
Matches 62; Conservative 33; Mismatches 102; Indels 46; Gaps 11;
```



```
Qy 9 VVLDGASKYLNEALWTFNDSPN---YYKTLGTSQITPALPPKAGDILYSKLDLGRTR 65
Db ||| ||| : ||| : : : : : ||| ||| ||| ||| |||
16 VVLCVGALQQNVEAKAKHYKTTSHVETQYVSTSSKKILP--FTHNKQIKVGPLDNLGRA- 72
Qy 66 TARGTLTYANVEGSYGVRSFGKNQ--NPAGWTGNPNHVKY-----KIEWLNGLSYVGD 117
Db ||| : : : ||| : : : ||| : : : ||| : : : |||
73 -----TYSHIQLRDADEPKIKRERLTYNPTGW----HNYKFTTEKGKTTWL----- 114
Qy 118 FWNRSHLIADSLGG-DALRVNAVGTGTRTQNVG-GRDQKGMRYTEQRAQEWLEANRDGYL 175
Db : ||| : : : ||| : : : ||| : : : ||| : : : |||
115 -MDRGHLVGYQFSGMNVNPNLVTMTKYLNTGFSENNPDGMLYYENRLDSWLANHKNFWL 173
Qy 176 YVEVAPIYNADLIPRAVVUSMQSSDNTINEKVLVYNTANG-----YTINYHNGTP 226
Db ||| ||| : ||| : : ||| : : ||| : : ||| : : |||
174 DYKVTPIYEGNNLPSRVELQYVGIDK--QGKLLLEIKLGGKEQTDYGVTTVTLENTSP 231
Qy 227 TQK 229
Db 232 LAK 234
```

Search completed: January 5, 2004, 18:41:36  
Job time : 55.1381 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:38:21 ; Search time 51.3149 Seconds  
(without alignments)  
762.478 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203

Sequence: 1 MNLGSRVFSKRLVKFS.....LVSATMVTVTLENTALAR 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.\*

1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/2/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/2/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/2/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/2/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	ID	Description
1	203	100.0	43	8 US-08-482-785-1 Sequence 1, Appli

2	203	100.0	43	15	US-09-119-900-1	Sequence 1, Appli
3	203	100.0	43	30	US-10-453-032-1	Sequence 1, Appli
4	203	100.0	271	5	US-08-188-721A-2	Sequence 2, Appli
5	203	100.0	271	8	US-08-482-785-8	Sequence 8, Appli
6	203	100.0	271	15	US-09-119-900-8	Sequence 8, Appli
7	203	100.0	271	30	US-10-453-032-8	Sequence 8, Appli
8	194	95.6	272	8	US-08-482-785-15	Sequence 15, Appli
9	194	95.6	272	15	US-09-119-900-15	Sequence 15, Appli
10	194	95.6	272	30	US-10-453-032-15	Sequence 15, Appli
11	59	29.1	219	30	US-10-437-963-121351	Sequence 121351,
12	58	28.6	308	18	US-09-489-039A-7674	Sequence 7674, Ap
13	58	28.6	308	30	US-10-446-203-7674	Sequence 7674, Ap
14	58	28.6	507	32	US-60-161-932-1687	Sequence 1687, Ap
15	57	28.1	216	22	US-09-758-472-6630	Sequence 6630, Ap
16	57	28.1	216	28	US-10-235-926-6630	Sequence 6630, Ap
17	57	28.1	1589	20	US-09-656-633-4	Sequence 4, Appli
18	57	28.1	1591	1	PCT-US01-00663-31088	Sequence 31088, A
19	57	28.1	1591	23	US-09-864-761-37952	Sequence 37952, A
20	57	28.1	1591	27	US-10-182-993-30159	Sequence 30159, A
21	57	28.1	1591	27	US-10-182-995-24424	Sequence 24424, A
22	57	28.1	1591	27	US-10-182-997-23157	Sequence 23157, A
23	57	28.1	1591	27	US-10-182-998-14678	Sequence 14678, A
24	57	28.1	1591	28	US-10-203-134-30801	Sequence 30801, A
25	57	28.1	1591	28	US-10-203-135-29801	Sequence 29801, A
26	57	28.1	1591	28	US-10-203-136-30825	Sequence 30825, A
27	57	28.1	1591	28	US-10-203-137-31088	Sequence 31088, A
28	57	28.1	1591	28	US-10-203-138-15081	Sequence 15081, A
29	57	28.1	1591	28	US-10-203-138A-15081	Sequence 15081, A
30	57	28.1	1591	28	US-10-203-139-30000	Sequence 30000, A
31	57	28.1	1591	32	US-60-236-359-20419	Sequence 20419, A
32	55.5	27.3	397	6	US-08-298-542-6	Sequence 6, Appli
33	55.5	27.3	397	7	US-08-330-275A-6	Sequence 6, Appli
34	55.5	27.3	397	17	US-09-300-597-20	Sequence 20, Appli
35	55.5	27.3	397	20	US-09-610-134-98	Sequence 98, Appli
36	55.5	27.3	397	22	US-09-791-537-79111	Sequence 79111, A
37	55.5	27.3	397	22	US-09-791-537-110035	Sequence 110035,
38	55.5	27.3	397	23	US-09-791-537-112028	Sequence 112028,
39	55.5	27.3	397	23	US-09-847-208-15	Sequence 15, Appli
40	55.5	27.3	397	23	US-09-847-208B-15	Sequence 15, Appli
41	55	27.1	341	22	US-09-791-537-4174	Sequence 4174, Ap
42	55	27.1	576	22	US-09-791-537-14717	Sequence 14717,
43	54	26.6	125	32	US-60-186-652-829	Sequence 829, App
44	54	26.6	927	28	US-10-289-762-472	Sequence 472, App
45	53	26.1	1471	20	US-09-614-150-3009	Sequence 3009, Ap

#### ALIGNMENTS

RESULT 1  
US-08-482-785-1  
; Sequence 1, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,785

```
;
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
US-08-482-785-1

Query Match 100.0%; Score 203; DB 8; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.5e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43

RESULT 2
US-09-119-900-1
; Sequence 1, Application US/09119900
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE: 23-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
US-10-453-032-1

Query Match 100.0%; Score 203; DB 30; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.5e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43

SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
```

```
;
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
US-09-119-900-1

Query Match 100.0%; Score 203; DB 15; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.5e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43

RESULT 3
US-10-453-032-1
; Sequence 1, Application US/10453032
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/453,032
; FILING DATE: 03-JUNE-2003
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE: 23-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
US-10-453-032-1

Query Match 100.0%; Score 203; DB 30; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.5e-22;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43
DB 1 MNLGSRVFSKCKRLVKFSWALVSATMAVTTVLTAL 43

SEQUENCE CHARACTERISTICS:
LENGTH: 43 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
```

Db 1 MNLGSRVFSKCRVKFSMVALVSATMAVTTVLTENTALAR 43

## RESULT 4

US-08-188-721A-2  
; Sequence 2, Application US/08188721A  
; GENERAL INFORMATION:  
; APPLICANT: Yutsudo, Takaashi  
; APPLICANT: Okumura, Koichi  
; APPLICANT: Iwasaki, Makoto  
; APPLICANT: Hara, Ayako  
; APPLICANT: Kishishita, Masamichi  
; APPLICANT: Takeda, Yoshifumi  
; APPLICANT: Igarashi, Hisanaga  
; APPLICANT: Hinuma, Yorio  
; TITLE OF INVENTION: Mitogenic Factor, Gene Thereof and  
; TITLE OF INVENTION: Method of Microdetection Therefor  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 8110 Gatehouse Road, Suite 500 East  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22042  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/188,721A  
; FILING DATE: 31-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Murthy Jr, Gerald M.  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 1422-178P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 205-8000  
; TELEFAX: (703) 205-8050  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-188-721A-2

Query Match 100.0%; Score 203; DB 5; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.6e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCRVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 1 MNLGSRVFSKCRVKFSMVALVSATMAVTTVLTENTALAR 43

## RESULT 5

US-08-482-785-8  
; Sequence 8, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena

; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,785  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE: 24-FEB-1995  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-785-8

Query Match 100.0%; Score 203; DB 8; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.6e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCRVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 1 MNLGSRVFSKCRVKFSMVALVSATMAVTTVLTENTALAR 43

## RESULT 6

US-09-119-900-8  
; Sequence 8, Application US/09119900  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-119-900-8

Query Match 100.0%; Score 203; DB 15; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.6e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 43

## RESULT 7

US-10-453-032-8

Sequence 8, Application US/10453032

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/453,032

FILING DATE: 03-JUNE-2003

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/082,845

FILING DATE: 23-JUNE-1993

ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.

REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 271 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-10-453-032-8

Query Match 100.0%; Score 203; DB 30; Length 271;  
Best Local Similarity 100.0%; Pred. No. 3.6e-21;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 43

## RESULT 8

US-08-482-785-15

Sequence 15, Application US/08482785

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/482,785

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/393,889

FILING DATE: 24-FEB-1995

APPLICATION NUMBER: US/08/082,845

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.

REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 272 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-482-785-15

Query Match 95.8%; Score 194; DB 8; Length 272;

Best Local Similarity 97.6%; Pred. No. 8e-20;

Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 4 NLLGSRVFSKCKRLVKFSMVALVSATMAVTTVLTENTALAR 45

## RESULT 9

US-09-119-900-15

Sequence 15, Application US/09119900

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
US-09-119-900-15

Query Match 95.6%; Score 194; DB 15; Length 272;  
Best Local Similarity 97.6%; Pred. No. 8e-20;  
Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NLLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 4 NLLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALAQ 45

RESULT 10  
US-10-453-032-15  
Sequence 15, Application US/10453032  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,032  
FILING DATE: 03-JUNE-2003  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: 23-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-10-453-032-15

Query Match 95.6%; Score 194; DB 30; Length 272;  
Best Local Similarity 97.6%; Pred. No. 8e-20;  
Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 NLLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALAR 43  
Db 4 NLLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTALAQ 45

RESULT 11  
US-10-437-963-121351  
Sequence 121351, Application US/10437963  
GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 121351  
LENGTH: 219  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(219)  
OTHER INFORMATION: unsure at all xaa locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_24385C.1.pep  
US-10-437-963-121351

Query Match 29.1%; Score 59; DB 30; Length 219;  
Best Local Similarity 41.0%; Pred. No. 8.5;  
Matches 16; Conservative 6; Mismatches 15; Indels 2; Gaps 1;

QY 2 NLLGSRVFSKKRLVKFSMVALVSATMAVTTVLTENTA 40  
Db 10 NLLNFQNTFMKPSL--FSVIALASAMAPAVQVQAKA 46

RESULT 12  
US-09-489-039A-7674  
Sequence 7674, Application US/09489039A  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709.2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 7674  
LENGTH: 308  
TYPE: PRT  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7674

Query Match 28.6%; Score 58; DB 18; Length 308;

Best Local Similarity 34.7%; Pred. No. 18;

Matches 17; Conservative 5; Mismatches 15; Indels 12; Gaps 1;

QY 5 GSRVFSKCKRLVKF-----SMVALVSATMAVTTVTLENTAL 41

Db 216 GLEQVFMDCRVAGTTPQVQRINDVTSMSLSVAGVGVVALVPLSARAL 264

RESULT 13

US-10-446-203-7674

; Sequence 7674, Application US/10446203

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/10/446.203

; PRIOR FILING DATE: 2003-05-27

; PRIOR APPLICATION NUMBER: US/09/489,039

; PRIOR FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 7674

; LENGTH: 308

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-10-446-203-7674

Query Match 28.6%; Score 58; DB 30; Length 308;

Best Local Similarity 34.7%; Pred. No. 18;

Matches 17; Conservative 5; Mismatches 15; Indels 12; Gaps 1;

QY 5 GSRVFSKCKRLVKF-----SMVALVSATMAVTTVTLENTAL 41

Db 216 GLEQVFMDCRVAGTTPQVQRINDVTSMSLSVAGVGVVALVPLSARAL 264

RESULT 14

US-60-161-932-1687

; Sequence 1687, Application US/60161932

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic

; FILE REFERENCE: CLO00122

; CURRENT APPLICATION NUMBER: US/60/161,932

; PRIOR FILING DATE: 1999-10-28

; NUMBER OF SEQ ID NOS: 2626

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1687

; LENGTH: 507

; TYPE: PRT

; ORGANISM: Drosophila

US-60-161-932-1687

Query Match 28.6%; Score 58; DB 32; Length 507;

Best Local Similarity 40.5%; Pred. No. 31;

Matches 17; Conservative 6; Mismatches 13; Indels 6; Gaps 2;

QY 2 NLLG-----SRRVFSKCKRLVKFSMVALVSATMAVTTVTLENT 39

Db 244 NLLSAQFKSEANFSKCLLL--TLTALVALSAAATTTTTTTT 283

RESULT 15

US-09-758-472-6630

; Sequence 6630, Application US/09758472

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

Query Match 28.1%; Score 57; DB 22; Length 216;

Best Local Similarity 34.3%; Pred. No. 17;

Matches 12; Conservative 13; Mismatches 8; Indels 2; Gaps 1;

QY 2 NLLGSRVFSKCKRLVKFSMVALVSATMAVTTVTTL 36

Db 50 NILKGNITKEAKL--FSFLALLNSYVPDITISL 82

Search completed: January 5, 2004, 18:55:04

Job time : 55.3149 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 5, 2004, 18:39:19 ; Search time 6.81031 Seconds  
(without alignments)  
408.499 Million cell updates/sec

Title: US-08-482-785-1

Perfect score: 203

Sequence: 1 MNLGSRVFSKKRLVKFS.....LVSATMAVTTVLTNTALAR 43

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 320769 seqs, 64697744 residues

Total number of hits satisfying chosen parameters: 320769

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	203	100.0	43	4	US-08-472-630-1
2	203	100.0	271	4	US-08-472-630-8
3	194	95.6	272	4	US-08-472-630-15
4	53	26.1	1471	5	US-09-614-150A-3009
5	52	25.6	304	6	US-10-425-114A-62081
6	52	25.6	336	6	US-10-679-063-6324
7	52	25.6	384	6	US-10-679-063-5446
8	52	25.6	384	6	US-10-679-063-6586
9	52	25.6	394	6	US-10-425-114A-50770
10	52	25.6	398	6	US-10-679-063-3812
11	51	25.1	730	5	US-09-614-150A-195
12	49	24.1	343	6	US-10-343-650A-660
13	49	24.1	426	6	US-10-650-274-31
14	48.5	23.9	458	6	US-10-425-114A-50060
15	48	23.6	335	6	US-10-425-114A-62521
16	48	23.6	3201	5	US-09-614-150A-15489
17	47.5	23.4	151	6	US-10-425-114A-44778
18	47	23.2	521	5	US-09-614-150A-2781
19	47	23.2	521	6	US-10-679-063-18636
20	47	23.2	1048	1	PCT-US03-38193-2687
21	47	23.2	1048	6	US-10-723-860-2687
22	47	23.2	1963	5	US-09-614-150A-15249
23	46.5	22.9	236	6	US-10-679-063-10401
24	46	22.7	163	6	US-10-679-063-2119
25	46	22.7	333	6	US-10-679-063-6285
26	46	22.7	889	6	US-10-679-063-12193

27	46	22.7	3432	6	US-10-679-063-14204	Sequence 14204, A
28	45.5	22.4	164	6	US-10-322-696A-132	Sequence 132, App
29	45.5	22.4	164	7	US-60-487-610-1573	Sequence 1573, Ap
30	45.5	22.4	164	7	US-60-485-450-1001	Sequence 1001, Ap
31	45.5	22.4	227	6	US-10-425-114A-61944	Sequence 61944, A
32	45.5	22.4	253	1	PCT-US03-32827-58	Sequence 58, Appl
33	45.5	22.4	253	6	US-10-655-873-8	Sequence 8, Appl
34	45.5	22.4	253	6	US-10-322-696A-186	Sequence 186, App
35	45.5	22.4	253	6	US-10-688-845-58	Sequence 58, Appl
36	45.5	22.4	253	7	US-60-487-610-1572	Sequence 1572, Ap
37	45.5	22.4	253	7	US-60-485-450-1000	Sequence 1000, Ap
38	45.5	22.4	288	5	US-09-897-516A-6710	Sequence 6710, Ap
39	45.5	22.4	488	6	US-10-679-063-12973	Sequence 12973, A
40	45	22.2	399	6	US-10-679-063-20248	Sequence 20248, A
41	45	22.2	413	6	US-10-296-115-1094	Sequence 1094, Ap
42	45	22.2	424	6	US-10-389-647-446	Sequence 446, App
43	45	22.2	520	6	US-10-679-063-5781	Sequence 5781, Ap
44	45	22.2	533	5	US-09-614-150A-18036	Sequence 18036, A
45	45	22.2	665	5	US-09-614-150A-30417	Sequence 30417, A

## ALIGNMENTS

### RESULT 1

US-08-472-630-1  
; Sequence 1, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Belsei, Marina  
; Pang, Patty P.-Y.  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-8321  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEetical: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-08-472-630-1

Query Match 100.0%; Score 203; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 2.2e-22;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCKRLVKFMSVALVSATMAVTTVTTLENTALAR 43  
Db 1 MNLGSRVFSKCKRLVKFMSVALVSATMAVTTVTTLENTALAR 43

RESULT 2  
US-08-472-630-8  
; Sequence 8, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:  
US-08-472-630-15  
Query Match 95.6%; Score 194; DB 4; Length 272;  
Best Local Similarity 97.6%; Pred. No. 2.9e-20;  
Matches 41; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLLGSRVFSKCKRLVKFMSVALVSATMAVTTVTTLENTALAR 43  
Db 4 NLLGSRVFSKCKRLVKFMSVALVSATMAVTTVTTLENTALAR 45

RESULT 4  
US-09-614-150A-3009  
; Sequence 3009, Application US/09614150A  
; GENERAL INFORMATION:  
; APPLICANT: Venter, J. Craig  
; APPLICANT: et al.  
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
; TITLE OF INVENTION: ARRAYS FOR DETECTING EXPRESSION OF 10,000 OR MORE  
; FILE REFERENCE: CL000728  
; CURRENT APPLICATION NUMBER: US/09/614,150A  
; CURRENT FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/157,832  
; PRIOR FILING DATE: 1999-10-05  
; PRIOR APPLICATION NUMBER: 60/160,191  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: 60/161,932  
; PRIOR FILING DATE: 1999-10-28  
; PRIOR APPLICATION NUMBER: 60/164,769  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/173,383  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/175,693  
; PRIOR FILING DATE: 2000-01-12  
; PRIOR APPLICATION NUMBER: 60/184,831  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: 60/191,637  
; PRIOR FILING DATE: 2000-03-23  
; NUMBER OF SEQ ID NOS: 43008

! SOFTWARE: FastSeq for Windows Version 4.0

! SEQ ID NO 3009

! LENGTH: 1471

! TYPE: PRT

! ORGANISM: DROSOPHILA

US-09-614-150A-3009

Query Match 26.1%; Score 53; DB 5; Length 1471;

Best Local Similarity 30.2%; Pred. No. 31;

Matches 13; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKKRLVKFSMVALVSATMAVTTVLTALAR 43

Db 76 MQLIGDRGKSLVLEVNARICATITTIQSSVEHMAER 118

RESULT 5

US-10-425-114A-62081

! Sequence 62081, Application US/10425114A

! GENERAL INFORMATION:

! APPLICANT: Liu, Jingdong

! APPLICANT: Zhou, Yihua

! APPLICANT: Kovalic, David K.

! APPLICANT: Screen, Steven E.

! APPLICANT: Tabaska, Jack E.

! APPLICANT: Cao, Yongwei

! TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

! FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

! FILE REFERENCE: 38-21(5313)B

! CURRENT APPLICATION NUMBER: US/10/425,114A

! CURRENT FILING DATE: 2003-04-28

! NUMBER OF SEQ ID NOS: 73128

! SEQ ID NO 62081

! LENGTH: 304

! TYPE: PRT

! ORGANISM: Zea mays

! FEATURE:

! OTHER INFORMATION: Clone ID: LIB3637-227-D4\_FLI.pep

US-10-425-114A-62081

Query Match 25.6%; Score 52; DB 6; Length 304;

Best Local Similarity 48.1%; Pred. No. 8.2;

Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 9 VFSKKRLVKFSMVALVSATMAVTTVT 35

Db 89 LFKKVSRLRKVSTLVVVSFGVAVATVT 115

RESULT 6

US-10-679-063-6324

! Sequence 6324, Application US/10679063

! GENERAL INFORMATION:

! APPLICANT: Edgerton, Michael D

! TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

! FILE REFERENCE: 38-15(52054)B

! CURRENT APPLICATION NUMBER: US/10/679,063

! CURRENT FILING DATE: 2003-10-02

! PRIOR APPLICATION NUMBER: 60/415,758

! PRIOR FILING DATE: 2002-10-02

! NUMBER OF SEQ ID NOS: 27373

! SEQ ID NO 6324

! LENGTH: 336

! TYPE: PRT

! ORGANISM: Zea mays

US-10-679-063-6324

Query Match 25.6%; Score 52; DB 6; Length 336;

Best Local Similarity 48.1%; Pred. No. 9.1;

Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 9 VFSKKRLVKFSMVALVSATMAVTTVT 35

Db 89 LFKKVSRLRKVSTLVVVSFGVAVATVT 115

Db 169 LFKKVSRLRKVSTLVVVSFGVAVATVT 195

RESULT 7

US-10-679-063-5446

! Sequence 5446, Application US/10679063

! GENERAL INFORMATION:

! APPLICANT: Edgerton, Michael D

! TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

! FILE REFERENCE: 38-15(52054)B

! CURRENT APPLICATION NUMBER: US/10/679,063

! CURRENT FILING DATE: 2003-10-02

! PRIOR APPLICATION NUMBER: 60/415,758

! PRIOR FILING DATE: 2002-10-02

! NUMBER OF SEQ ID NOS: 27373

! SEQ ID NO 5446

! LENGTH: 384

! TYPE: PRT

! ORGANISM: Zea mays

US-10-679-063-5446

Query Match 25.6%; Score 52; DB 6; Length 384;

Best Local Similarity 48.1%; Pred. No. 10;

Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 9 VFSKKRLVKFSMVALVSATMAVTTVT 35

Db 169 LFKKVSRLRKVSTLVVVSFGVAVATVT 195

RESULT 8

US-10-679-063-6586

! Sequence 6586, Application US/10679063

! GENERAL INFORMATION:

! APPLICANT: Edgerton, Michael D

! TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES

! FILE REFERENCE: 38-15(52054)B

! CURRENT APPLICATION NUMBER: US/10/679,063

! CURRENT FILING DATE: 2003-10-02

! PRIOR APPLICATION NUMBER: 60/415,758

! PRIOR FILING DATE: 2002-10-02

! NUMBER OF SEQ ID NOS: 27373

! SEQ ID NO 6586

! LENGTH: 384

! TYPE: PRT

! ORGANISM: Zea mays

US-10-679-063-6586

Query Match 25.6%; Score 52; DB 6; Length 384;

Best Local Similarity 48.1%; Pred. No. 10;

Matches 13; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 9 VFSKKRLVKFSMVALVSATMAVTTVT 35

Db 169 LFKKVSRLRKVSTLVVVSFGVAVATVT 195

RESULT 9

US-10-425-114A-50770

! Sequence 50770, Application US/10425114A

! GENERAL INFORMATION:

! APPLICANT: Liu, Jingdong

! APPLICANT: Zhou, Yihua

! APPLICANT: Kovalic, David K.

! APPLICANT: Screen, Steven E.

! APPLICANT: Tabaska, Jack E.

! APPLICANT: Cao, Yongwei

! TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

! FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement

! FILE REFERENCE: 38-21(5313)B

! CURRENT APPLICATION NUMBER: US/10/425,114A

! CURRENT FILING DATE: 2003-04-28

! NUMBER OF SEQ ID NOS: 73128



Search completed: January 5, 2004, 18:56:43  
Job time : 9.81031 secs

Best Local Similarity 37.2%; Pred. No. 32;  
Matches 16; Conservative 5; Mismatches 12; Indels 10; Gaps 2;

Qy 5 GSRVFS-----KKRLVKFSVALVSATMAVTTVLENT 39  
Db 370 GSARFFSPLGVYDFVKRTQFIQYSAAAL--ATQADAIVTLAQT 410

## RESULT 14

US-10-425-114A-50060  
; Sequence 50060, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 50060  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3592-069-B11\_FLI.pep  
US-10-425-114A-50060

Query Match 23.9%; Score 48.5; DB 6; Length 458;  
Best Local Similarity 37.1%; Pred. No. 40;  
Matches 13; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

Qy 7 RRVFSKRCRL-VKFSVALVSATMAVTTVLENTA 40  
Db 1 RRMRTTRCMLMVAQALVVVVSALLAATATTTTAA 35

## RESULT 15

US-10-425-114A-62521  
; Sequence 62521, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 62521  
; LENGTH: 335  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700423405\_FLI.pep  
US-10-425-114A-62521

Query Match 23.6%; Score 48; DB 6; Length 335;  
Best Local Similarity 35.1%; Pred. No. 34;  
Matches 13; Conservative 8; Mismatches 14; Indels 2; Gaps 1;

Qy 8 RVFSKRCRLVKF--SMVALVSATMAVTTVLENTALA 42  
Db 194 RYFEKOWDMVSFIRSLKAMVRSSNAVAVTFPSTVLS 230

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:53:52 ; Search time 3240 Seconds  
(without alignments)  
11085.300 Million cell updates/sec

Title: US-08-482-785-7  
Perfect score: 1083  
Sequence: 1 GACAACGCCCTCTCTTTTCT.....TGTGCAAAAGCAAAAGC 1083

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 33363688 seqs, 16581889874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA Main:  
1: /cgn2\_6/ptodata/2/pna/PCTUS COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/PCTUS COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US06 COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US07 COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US08 COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
29: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US09 COMB.seq.\*

44: /cgn2\_6/ptodata/2/pna/US100A COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US100B COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US101A COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US101B COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US102A COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US102B COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US103A COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US103B COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US104A COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US104B COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US6000 COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US6001 COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6002 COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6003 COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6004 COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6005 COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6006 COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6007 COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6008 COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6009 COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6010 COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6011 COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6012 COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6013 COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6014 COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6015 COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6016 COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6017 COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6018 COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6019 COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6020 COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6021 COMB.seq.\*  
76: /cgn2\_6/ptodata/2/pna/US6022 COMB.seq.\*  
77: /cgn2\_6/ptodata/2/pna/US6023 COMB.seq.\*  
78: /cgn2\_6/ptodata/2/pna/US6023B COMB.seq.\*  
79: /cgn2\_6/ptodata/2/pna/US6024 COMB.seq.\*  
80: /cgn2\_6/ptodata/2/pna/US6025 COMB.seq.\*  
81: /cgn2\_6/ptodata/2/pna/US6026 COMB.seq.\*  
82: /cgn2\_6/ptodata/2/pna/US6027 COMB.seq.\*  
83: /cgn2\_6/ptodata/2/pna/US6028 COMB.seq.\*  
84: /cgn2\_6/ptodata/2/pna/US6029 COMB.seq.\*  
85: /cgn2\_6/ptodata/2/pna/US6030 COMB.seq.\*  
86: /cgn2\_6/ptodata/2/pna/US6031 COMB.seq.\*  
87: /cgn2\_6/ptodata/2/pna/US6032 COMB.seq.\*  
88: /cgn2\_6/ptodata/2/pna/US6033 COMB.seq.\*  
89: /cgn2\_6/ptodata/2/pna/US6034 COMB.seq.\*  
90: /cgn2\_6/ptodata/2/pna/US6035 COMB.seq.\*  
91: /cgn2\_6/ptodata/2/pna/US6036 COMB.seq.\*  
92: /cgn2\_6/ptodata/2/pna/US6037 COMB.seq.\*  
93: /cgn2\_6/ptodata/2/pna/US6038 COMB.seq.\*  
94: /cgn2\_6/ptodata/2/pna/US6039 COMB.seq.\*  
95: /cgn2\_6/ptodata/2/pna/US6040 COMB.seq.\*  
96: /cgn2\_6/ptodata/2/pna/US6041 COMB.seq.\*  
97: /cgn2\_6/ptodata/2/pna/US6042 COMB.seq.\*  
98: /cgn2\_6/ptodata/2/pna/US6043 COMB.seq.\*  
99: /cgn2\_6/ptodata/2/pna/US6044 COMB.seq.\*  
100: /cgn2\_6/ptodata/2/pna/US6045 COMB.seq.\*  
101: /cgn2\_6/ptodata/2/pna/US6046 COMB.seq.\*  
102: /cgn2\_6/ptodata/2/pna/US6047 COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	DB ID	Description
1	1083	100.0	1083	9 US-08-482-785-7 Sequence 7, Appli
2	1083	100.0	1083	16 US-09-119-900-7 Sequence 7, Appli
3	1083	100.0	1083	53 US-10-453-032-7 Sequence 7, Appli
4	961.6	88.8	1021	6 US-08-188-721-1 Sequence 1, Appli





Db 601 GAGATTCTCGAATAGAGTCATCTCTATTCAGATAGTCTCGGTGGAGATGCACTCAGAG 660  
Qy 661 TCAATGCGGTACAGGAACAGTACCCAAATGTAGGAGTCTGTGACCAAAAGGCGCA 720  
Db 661 TCAATGCGGTACAGGAACAGTACCCAAATGTAGGAGTCTGTGACCAAAAGGCGCA 720  
Qy 721 TGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Db 721 TGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Qy 781 ATTATGAAGTCGCTCCCAATCTCAACGACGAGTGTATTCAGAGCTGTCTGTAT 840  
Db 781 ATTATGAAGTCGCTCCCAATCTCAACGACGAGTGTATTCAGAGCTGTCTGTAT 840  
Qy 841 CAATGCAATCTCTGATAATACCACTCAACGAGAAAGTATTAGTTTACCAACAGCTAATG 900  
Db 841 CAATGCAATCTCTGATAATACCACTCAACGAGAAAGTATTAGTTTACCAACAGCTAATG 900  
Qy 901 GCTACACCATTAACCTACCATTAACGCTACCTCAAAATTAATACCAAAAGGCTAGAC 960  
Db 901 GCTACACCATTAACCTACCATTAACGCTACCTCAAAATTAATACCAAAAGGCTAGAC 960  
Qy 961 CTCTGCTCACTAGGCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAA 1020  
Db 961 CTCTGCTCACTAGGCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAA 1020  
Qy 1021 TACTAGAAAAAGCAATGATTCGCTCAATGCTTTTATGAATTTGTCAAAAGCAAAAA 1080  
Db 1021 TACTAGAAAAAGCAATGATTCGCTCAATGCTTTTATGAATTTGTCAAAAGCAAAAA 1080  
Qy 1081 AGC 1083  
Db 1081 AGC 1083

## RESULT 2

US-09-119-900-7  
; Sequence 7, Application US/09119900  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patti P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1083 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 129..944  
; US-09-119-900-7

Query Match 100.0%; Score 1083; DB 16; Length 1083;

Best Local Similarity 100.0%; Pred. No. 3e-239;

Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAAAGCGCTTCTTTTCTCTCTTACTATCTCTCTTAATTTTTCATATTTTAAAAAAC 60  
Db 1 GACAAAGCGCTTCTTTTCTCTCTTACTATCTCTCTTAAATTTTTCATATTTTAAAAAAC 60  
Qy 61 TATTGATAAACTAGTTTAAAGTAAAGCGGTATCTATGGTTAGTAGCGAAATTAGAAAGAGG 120  
Db 61 TATTGATAAACTAGTTTAAAGTAAAGCGGTATCTATGGTTAGTAGCGAAATTAGAAAGAGG 120  
Qy 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180  
Db 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180  
Qy 181 AATTTTCAATGGTAGTCTTTGTATCAGCCCAATGCTGTAAACAGTACACACTTGA 240  
Db 181 AATTTTCAATGGTAGTCTTTGTATCAGCCCAATGCTGTAAACAGTACACACTTGA 240  
Qy 241 ATACTGCACCTGGCAGCACCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCGCAA 300  
Db 241 ATACTGCACCTGGCAGCACCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCGCAA 300  
Qy 301 GCAAGTACCTAAACGAAAGCAATAGCTTTGGACATTTCAATGACAGTCTTAACTATTACAAA 360  
Db 301 GCAAGTACCTAAACGAAAGCAATAGCTTTGGACATTTCAATGACAGTCTTAACTATTACAAA 360  
Qy 361 CTTTAGTACTAGTCAAGTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTTATA 420  
Db 361 CTTTAGTACTAGTCAAGTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTTATA 420  
Qy 421 GCAATTTAGATGATTTAGGAGGACGCTACTCTGAGAGTACATTCACCTTATGCCAATG 480  
Db 421 GCAATTTAGATGATTTAGGAGGACGCTACTCTGAGAGTACATTCACCTTATGCCAATG 480  
Qy 481 TTGAAGGTAGCTACGGTGTAGACAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGGA 540  
Db 481 TTGAAGGTAGCTACGGTGTAGACAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGGA 540  
Qy 541 CTGGAACCCCTAATCATGTCAAAATATAAAATGAATGGTTAAATGGTCTATCTTTATGTCG 600  
Db 541 CTGGAACCCCTAATCATGTCAAAATATAAAATGAATGGTTAAATGGTCTATCTTTATGTCG 600  
Qy 601 GAGATTTCTGGAATAGAGTCAATCTCAATTCAGATAGTCTCGGTGGAGATGCACTCAGAG 660  
Db 601 GAGATTTCTGGAATAGAGTCAATCTCAATTCAGATAGTCTCGGTGGAGATGCACTCAGAG 660  
Qy 661 TCAATGCCGTTTACAGGAACACGCTACCCAAATGTAGGAGTCTGTGACCAAAAGGCGGCA 720  
Db 661 TCAATGCCGTTTACAGGAACACGCTACCCAAATGTAGGAGTCTGTGACCAAAAGGCGGCA 720  
Qy 721 TGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Db 721 TGGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Qy 781 ATTATGAAGTCGCTCCCAATCTCAACGACGAGTGTATTCAGAGCTGTCTGTAT 840  
Db 781 ATTATGAAGTCGCTCCCAATCTCAACGACGAGTGTATTCAGAGCTGTCTGTAT 840

Qy	841	CAATGCAATCTTCTGATAATACGATCAACGAGAAAGTATTAGTTTACAACACAGCTTAATG	900
Db	841	CAATGCAATCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTACAACACAGCTTAATG	900
Qy	901	GCTACACCATTAACCTACTACCATAAACGGTACACCTACTCAAAAATAATACCAAAAGGCTAGAC	960
Db	901	GCTACACCATTAACCTACTACCATAAACGGTACACCTACTCAAAAATAATACCAAAAGGCTAGAC	960
Qy	961	CTCTGCTCACTAGCGCTAGCTTTTTCATCAAAAAAGCAATGACTATAGAAAGTAAAAA	1020
Db	961	CTCTGCTCACTAGCGCTAGCTTTTTCATCAAAAAAGCAATGACTATAGAAAGTAAAAA	1020
Qy	1021	TACTAGAAAAAGCAATGATTCGCGCTCATTCGTTTTATGAATTTGTCGCAAAAGCAAAAA	1080
Db	1021	TACTAGAAAAAGCAATGATTCGCGCTCATTCGTTTTATGAATTTGTCGCAAAAGCAAAAA	1080
Qy	1081	AGC 1083	
Db	1081	AGC 1083	

### RESULT 3

```

US-10-453-032-7
: Sequence 7, Application US/10453032
: GENERAL INFORMATION:
: APPLICANT: Adams, Craig W.
: APPLICANT: Pang, Patty P.-Y.
: APPLICANT: Belei, Marina
: TITLE OF INVENTION: Recombinant DNase B Derived from
: TITLE OF INVENTION: Streptococcus pyogenes
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sheldon & Mak
: STREET: 225 South Lake Avenue, Ninth Floor
: CITY: Pasadena
: STATE: California
: COUNTRY: USA
: ZIP: 91001
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/10/453,032
: FILING DATE: 03-JUNE-2003
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/082,845
: FILING DATE: 23-JUNE-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Farber, Michael B.
: REGISTRATION NUMBER: 32,612
: REFERENCE/DOCKET NUMBER: 9521
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (818) 796-4000
: TELEFAX: (818) 795-6321
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1083 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Streptococcus pyogenes
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 129..944
: US-10-453-032-7

```

Qy 1021 TACTAGAAAAAGCAATGATTCCCGTCATTGCTTTTATGATTTGTCAAAAACCAAAA 1080  
Db |||||  
Qy 1021 TACTAGAAAAAGCAATGATTCCCGTCATTGCTTTTATGATTTGTCAAAAACCAAAA 1080  
Db |||||

Qy 1081 AGC 1083  
Db |||||

1081 AGC 1083

RESULT 4

US-08-188-721-1

; Sequence 1, Application US/08188721

; GENERAL INFORMATION:

; APPLICANT: Yutsudo, Takashi

; APPLICANT: Okumura, Koichi

; APPLICANT: Iwasaki, Makoto

; APPLICANT: Hara, Ayako

; APPLICANT: Kishishita, Masamichi

; APPLICANT: Takeda, Yoshifumi

; APPLICANT: Igarashi, Hisanaga

; APPLICANT: Hinuma, Yorio

; TITLE OF INVENTION: Mitogenic Factor, Gene Thereof and

; TITLE OF INVENTION: Method of Microdetection Therefor

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 8110 Gatehouse Road, Suite 500 East

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: U.S.A.

; ZIP: 22042

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/08/188, 721

; APPLICATION NUMBER: US/08/188, 721

; FILING DATE: 31-JAN-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Murphy Jr., Gerald M.

; REGISTRATION NUMBER: 28,977

; REFERENCE/DOCKET NUMBER: 1422-178P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 205-8000

; TELEFAX: (703) 205-8050

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1021 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pyogenes

; STRAIN: NY-5

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 155..967

US-08-188-721-1

Query Match 88.8%; Score 961.6; DB 6; Length 1021;  
Best Local Similarity 98.5%; Pred. No. 3.1e-211;  
Matches 981; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

Qy 1 GACACGCCCTCTCTTTTCTCCCTTACTATCTCTCTTAAATTTTCATATTTTAAATAAAC 60  
Db |||||

27 GACAAATGCCCTCTCTTTTCTCCCTTATATCTCGTTAAATTTTCATATTTTAAATAAAC 86

Qy 61 TATTGATAAACTAGTTAAAGTAAGCGGTATACCTATGTTAGTTAGCGAAATTTAGAAAAGG 120  
Db |||||

87 TATTGATAAACTAGTTAAAGTAAGCGGTATACCTATGTTAGTTAGCGAAATTTAGAAAAGG 146  
Qy 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAA 180  
Db |||||

147 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTTCTAAAAAATGTCGGCTAGTAA 206  
Qy 181 AATTTTCAATGGTAGTCTTTGTTATAGCCCAATAGGCTGTAAACAACAGTACACTTGAAA 240  
Db |||||

207 AATTTTCAATGGTAGTCTTTGTTATAGCCCAATAGGCTGTAAACAACAGTACACTTGAAA 266  
Qy 241 ATACTGCACCTGGCAGCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCCAA 300  
Db |||||

267 ATACTGCACCTGGCAGCAACACAGCTCTCAATGATGTTGTTCTAAATGATGGCCAA 326  
Qy 301 GCAAGTACCTAAACGAAGCAATTTAGCTTGGACATTTCAATGACAGTCTTAATTTACAAA 360  
Db |||||

327 CGAAGTACCTAAACGAAGCAATTTAGCTTGGACATTTCAATGACAGTCTTAATTTACAAA 386  
Qy 361 CTTTAGGTACTAGTCAAGTACTCCAGCACTCTTTCTAAAGCAGAGATATTTCTTATA 420  
Db |||||

387 CTTTAGGTACTAGTCAAGTACTCCAGCACTCTTTCTAAAGCAGAGATATTTCTTATA 446  
Qy 421 GCAATTTAGTAGTGTAGGAGACCGGTACTCTGCTAGAGGTACATTTGCACTTATGCCAATG 480  
Db |||||

447 GCAATTTAGTAGTGTAGGAGACCGGTACTCTGCTAGAGGTACATTTGCACTTATGCCAATG 506  
Qy 481 TTGAAGGTAGTACGGTGTGTAGACAATCTTTCCGTTAAATCAAAACCCCGCAGGATGGA 540  
Db |||||

507 TTGAAGGTAGTACGGTGTGTAGACAATCTTTCCGTTAAATCAAAACCCCGCAGGATGGA 566  
Qy 541 CTGGAACCCCTAATCATGTCAAAATATAAAATTTGAATGGTTAAATTTGCTTATCTTATGCG 600  
Db |||||

567 CTGGAACCCCTAATCATGTCAAAATATAAAATTTGAATGGTTAAATTTGCTTATCTGCG 626  
Qy 601 GAGATTTCTGGAATAGAGTCATCTCATTTGCAGATAGTCTCGGTGGAGATGCATCTCAGAG 660  
Db |||||

627 GAGATTTCTGGAATAGAGTCATCTCATTTGCAGATAGTCTCGGTGGAGATGCATCTCAGAG 686  
Qy 661 TCAATGCCGTTACAGGACACGCTACCCAAATCTAGGAGTCTGTCACCAAAAGGCGGCA 720  
Db |||||

687 TCAATGCCGTTACAGGACACGCTACCCAAATCTAGGAGTCTGTCACCAAAAGGCGGCA 746  
Qy 721 TGGCTATATACCGAACAAAGAGCTCAAGAAATGGTTAGAGCAAAATCGTGATGGCTATCTTT 780  
Db |||||

747 TGGCTATATACCGAACAAAGAGCTCAAGAAATGGTTAGAGCAAAATCGTGATGGCTATCTTT 806  
Qy 781 ATTATGAAGTGGCTCCAAATCTACAAACGACAGCGAGATTTCAAGAGTGTGCTGGTAT 840  
Db |||||

807 ATTATGAAGCGGCTCCAAATCTACAAACGACAGCGAGATTTCAAGAGTGTGCTGGTAT 866  
Qy 841 CAATGCAATCTTCTGATATATACCATCAACGAGAAAGTATTAGTTTACAAACAGCTAATG 900  
Db |||||

867 CAATGCAATCTTCTGATATATATCAACGAGAAAGTATTAGTTTACAAACAGCTAATG 926  
Qy 901 GCTACACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 960  
Db |||||

927 GCTACACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 986

961 CTCTGCTCACTAGGCTAGCTTTTATACATCAAAAAA 996  
Db |||||

987 CTCTGCTCACTAGG-CTAGCTTTTATACATCAAAAAA 1021

RESULT 5

US-08-188-721A-1

; Sequence 1, Application US/08188721A

; GENERAL INFORMATION:

; APPLICANT: Yutsudo, Takashi

; APPLICANT: Okumura, Koichi

; APPLICANT: Iwasaki, Makoto

; APPLICANT: Hara, Ayako



;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/482,785  
;; FILING DATE: 07-JUN-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/393,889  
;; FILING DATE: 24-FEB-1995  
;; APPLICATION NUMBER: US/08/082,845  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Farber, Michael B.  
;; REGISTRATION NUMBER: 32,612  
;; REFERENCE/DOCKET NUMBER: 9521  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (818) 795-4000  
;; TELEFAX: (818) 795-6321  
;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 940 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Streptococcus pyogenes  
;;  
US-08-482-785-11

Query Match 85.2%; Score 922.6; DB 9; Length 940;  
Best Local Similarity 99.6%; Pred. No. 3e-202;  
Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 123 AAGCATATGAATCTACTTGATCAAGACGGGTTTTTCTTAAATAATGTGGCTAGTAAAA 182  
DB 1 ATGATCCGAATCTACTTGATCAAGACGGGTTTTTCTTAAATAATGTGGCTAGTAAAA 60  
QY 183 TTTTCAATGTAGTCTTTGATCAGGCACAAATGGCTGTAAACAGTCACTTGAATAAT 242  
DB 61 TTTTCAATGTAGTCTTTGATCAGGCACAAATGGCTGTAAACAGTCACTTGAATAAT 120  
QY 243 ACTGCACTGGCAGCAAAACACAGGCTCTCAAAATGATGTTGTTCTTAATGATGGCGAAGC 302  
DB 121 ACTGCACTGGCAGCAAAACACAGGCTCTCAAAATGATGTTGTTCTTAATGATGGCGAAGC 180  
QY 303 AAGTACTTAAACGAAGCATTAGCTTGACATTCATCAAGTCACTTAACATTAACAAACT 362  
DB 181 AAGTACTTAAACGAAGCATTAGCTTGACATTCATCAAGTCACTTAACATTAACAAACT 240  
QY 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTTATAGC 422  
DB 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTTATAGC 300  
QY 423 AATTAGATGAGTTAGGAGGACGCTAGCTAGAGGTACATTTGATGATGATGATGATGATGAT 482  
DB 301 AATTAGATGAGTTAGGAGGACGCTAGCTAGAGGTACATTTGATGATGATGATGATGATGATG 360  
QY 483 GAAGGTAGCTACGGTGTAGACAATCTTTCCGTTAAATAATCAAAACCCCGCAGGATGACT 542  
DB 361 GAAGGTAGCTACGGTGTAGACAATCTTTCCGTTAAATAATCAAAACCCCGCAGGATGACT 420  
QY 543 GGAACCCCTAATCATGTCAAAATATAAAATGGAATGTTAAATGGTCTATCTTTATGTCGGA 602  
DB 421 GGAACCCCTAATCATGTCAAAATATAAAATGGAATGTTAAATGGTCTATCTTTATGTCGGA 480  
QY 603 GATTTCTGGAATAGAGTCACTCATTCGATGATGATGATGATGATGATGATGATGATGATGATG 662  
DB 481 GATTTCTGGAATAGAGTCACTCATTCGATGATGATGATGATGATGATGATGATGATGATGATG 540  
QY 663 AATGCGGTACAGGAACAGTCAACCAATGTAGGAGGTGCTGACCAAAAGGCGGCATG 722  
DB 541 AATGCGGTACAGGAACAGTCAACCAATGTAGGAGGTGCTGACCAAAAGGCGGCATG 600

QY 723 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 782  
DB 601 CGCTATACCGAACAAAGAGCTCAAGATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 660  
QY 783 TATGAAGTCGCTCCAATCTTACAAACGAGACGAGTGTGATTTCAAGAGAGCTGTGCTGGTATCA 842  
DB 661 TATGAAGTCGCTCCAATCTTACAAACGAGACGAGTGTGATTTCAAGAGAGCTGTGCTGGTATCA 720  
QY 843 ATGCAATCTTCTGATTAATACCAATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGGC 902  
DB 721 ATGCAATCTTCTGATTAATACCAATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGGC 780  
QY 903 TACACCATTAACCTACCAATCAACGAGTCAACCTACTCAAAATAATAACCAAAAGGCTAGACCT 962  
DB 781 TACACCATTAACCTACCAATCAACGAGTCAACCTACTCAAAATAATAACCAAAAGGCTAGACCT 840  
QY 963 CTGCTCACTAGGCTAGCTATGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAATA 1022  
DB 841 CTGCTCACTAGGCTAGCTATGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAATA 900  
QY 1023 CTAGAAAAAGCAATGATGCTCGCTCATGTC 1051  
DB 901 CTAGAAAAAGCAATGATGCTCGCTCATGTC 929

RESULT 7  
US-09-119-900-11  
;; Sequence 11, Application US/09119900  
;; GENERAL INFORMATION:  
;; APPLICANT: Adams, Craig W.  
;; APPLICANT: Pang, Patty P.-Y.  
;; APPLICANT: Belei, Marina  
;; TITLE OF INVENTION: Recombinant DNase B Derived from  
;; TITLE OF INVENTION: Streptococcus pyogenes  
;; NUMBER OF SEQUENCES: 16  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sheldon & Mak  
;; STREET: 225 South Lake Avenue, Ninth Floor  
;; CITY: Pasadena  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 91001  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION NUMBER: US/09/119,900  
;; FILING DATE:  
;; APPLICATION DATA:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Farber, Michael B.  
;; REGISTRATION NUMBER: 32,612  
;; REFERENCE/DOCKET NUMBER: 9521  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (818) 795-4000  
;; TELEFAX: (818) 795-6321  
;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 940 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;; ORIGINAL SOURCE:  
;; ORGANISM: Streptococcus pyogenes  
;;  
US-09-119-900-11

Query Match 85.2%; Score 922.6; DB 16; Length 940;  
Best Local Similarity 99.6%; Pred. No. 38-202;  
Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATCGGCTAGTAAAA 182  
Db 1 ATGATCCGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATCGGCTAGTAAAA 60

Qy 183 TTTTCAATGGTAGCTCTTGTATCAGCCACAATGGCTGTAAACAAGTCACTTGAATA 242  
Db 61 TTTTCAATGGTAGCTCTTGTATCAGCCACAATGGCTGTAAACAAGTCACTTGAATA 120

Qy 243 ACTGCATGCGACGACAAACACAGGCTCAATGATGTTCTTAAATGATGGCGAAGC 302  
Db 121 ACTGCATGCGACGACAAACACAGGCTCAATGATGTTCTTAAATGATGGCGAAGC 180

Qy 303 AAGTACCTAAACGAGCATTAGCTTGGACATTCATGACAGTCTTAATACTATTCAAAACT 362  
Db 181 AAGTACCTAAACGAGCATTAGCTTGGACATTCATGACAGTCTTAATACTATTCAAAACT 240

Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTCTTAAAGCAGGAGATTTCTCTATAGC 422  
Db 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTCTTAAAGCAGGAGATTTCTCTATAGC 300

Qy 423 AAATTAGATGATTAGGAAGGACCGTACTGCTAGAGTACATTTGACTTTATGCCAATGTT 482  
Db 301 AAATTAGATGATTAGGAAGGACCGTACTGCTAGAGTACATTTGACTTTATGCCAATGTT 360

Qy 483 GAAGGTAGTACGGTGTAGACAACTCTTTCGTAATAAATAAATCAAAACCGCAGGATGACT 542  
Db 361 GAAGGTAGTACGGTGTAGACAACTCTTTCGTAATAAATAAATCAAAACCGCAGGATGACT 420

Qy 543 GGAAACCTTAATCATGCAATATAAATTAATGATGTTAAATGCTCTATCTTATGTCGA 602  
Db 421 GGAAACCTTAATCATGCAATATAAATTAATGATGTTAAATGCTCTATCTTATGTCGA 480

Qy 603 GATTCTCGATAGAGTCACTCTCATTCGACATAGTCTCGGTGAGATGCACTCAGAGTC 662  
Db 481 GATTCTCGATAGAGTCACTCTCATTCGACATAGTCTCGGTGAGATGCACTCAGAGTC 540

Qy 663 AATCGGTTACAGAAACACGATACCAAAATGTAGAGGTCGTGACCAAAAGGCGGATG 722  
Db 541 AATCGGTTACAGAAACACGATACCAAAATGTAGAGGTCGTGACCAAAAGGCGGATG 600

Qy 723 CGTATACCGAAACAAAGAGTCAAGATGTTAGAGCAAAATCGTATGCTATCTTTAT 782  
Db 601 CGTATACCGAAACAAAGAGTCAAGATGTTAGAGCAAAATCGTATGCTATCTTTAT 660

Qy 783 TATGAAGTCGCTCCAAATCTACACGACAGGTTGATTCACAGAGTCTGCTGATCA 842  
Db 661 TATGAAGTCGCTCCAAATCTACACGACAGGTTGATTCACAGAGTCTGCTGATCA 720

Qy 843 ATGCAATCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTAAACACAGCTTAATGCG 902  
Db 721 ATGCAATCTTCTGATAATACCATCAACGAGAAAGTATTAGTTTAAACACAGCTTAATGCG 780

Qy 903 TACACATTAATACCATTAACGATACCTTCTTAAATAATATACCAAAAGGCTAGACCT 962  
Db 781 TACACATTAATACCATTAACGATACCTTCTTAAATAATATACCAAAAGGCTAGACCT 840

Qy 963 CTGCTCACTAGGCTAGCTTTTATCATCAAAAGCAATGACTATAGAAAGTAAATA 1022  
Db 841 CTGCTCACTAGGCTAGCTTTTATCATCAAAAGCAATGACTATAGAAAGTAAATA 900

Qy 1023 CTAGAAAAGCAATGATTGCGCTATTGC 1051  
Db 901 CTAGAAAAGCAATGATTGCGCTATTGC 929

RESULT 8  
US-10-453-032-11  
; Sequence 11, Application US/10453032

GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,032  
FILING DATE: 03-JUNE-2003  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: 23-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 940 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-10-453-032-11

Query Match 85.2%; Score 922.6; DB 53; Length 940;  
Best Local Similarity 99.6%; Pred. No. 38-202;  
Matches 925; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATCGGCTAGTAAAA 182  
Db 1 ATGATCCGAATCTACTTGGATCAAGACGGGTTTTCTTAAATAATCGGCTAGTAAAA 60

Qy 183 TTTTCAATGGTAGCTCTTGTATCAGCCACAATGGCTGTAAACAAGTCACTTGAATA 242  
Db 61 TTTTCAATGGTAGCTCTTGTATCAGCCACAATGGCTGTAAACAAGTCACTTGAATA 120

Qy 243 ACTGCATGCGACGACAAACACAGGCTCAATGATGTTCTTAAATGATGGCGAAGC 302  
Db 121 ACTGCATGCGACGACAAACACAGGCTCAATGATGTTCTTAAATGATGGCGAAGC 180

Qy 303 AAGTACCTAAACGAGCATTAGCTTGGACATTCATGACAGTCTTAATACTATTCAAAACT 362  
Db 181 AAGTACCTAAACGAGCATTAGCTTGGACATTCATGACAGTCTTAATACTATTCAAAACT 240

Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTCTTAAAGCAGGAGATTTCTCTATAGC 422  
Db 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTCTTAAAGCAGGAGATTTCTCTATAGC 300

Qy 423 AAATTAGATGATTAGGAAGGACCGTACTGCTAGAGTACATTTGACTTTATGCCAATGTT 482  
Db 301 AAATTAGATGATTAGGAAGGACCGTACTGCTAGAGTACATTTGACTTTATGCCAATGTT 360

Qy 483 GAAGGTAGCTACGGTGTGTAGACAATCTTTCGGTAAATAATCAAAACCCCGCAGGATGGACT 542  
Db |||||||  
Qy 361 GAAGGTAGCTACGGTGTGTAGACAATCTTTCGGTAAATAATCAAAACCCCGCAGGATGGACT 420  
Db |||||||  
Qy 543 GGAACCCCTAATCATGTCAAAATATAAAATTCGAATGGTTAAATGCTCTATCTTATGTCGGA 602  
Db |||||||  
Qy 421 GGAACCCCTAATCATGTCAAAATATAAAATTCGAATGGTTAAATGCTCTATCTTATGTCGGA 480  
Db |||||||  
Qy 603 GATTTCGGAATAGAAGTCATCTTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 662  
Db |||||||  
Qy 481 GATTTCGGAATAGAAGTCATCTTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 540  
Db |||||||  
Qy 663 AATCCGTTACAGAAACAGTCACCAAAATGTAGAGTCTGTGACCAAAAGGCGGCATG 722  
Db |||||||  
Qy 541 AATCCGTTACAGAAACAGTCACCAAAATGTAGAGTCTGTGACCAAAAGGCGGCATG 600  
Db |||||||  
Qy 723 CGCTATACCGAAACAAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 782  
Db |||||||  
Qy 601 CGCTATACCGAAACAAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 660  
Db |||||||  
Qy 783 TATGAAGTCGCTCCAAATCTACAACGACAGAGTGTGATTCGAAGAGCTGTGCTGGTATCA 842  
Db |||||||  
Qy 661 TATGAAGTCGCTCCAAATCTACAACGACAGAGTGTGATTCGAAGAGCTGTGCTGGTATCA 720  
Db |||||||  
Qy 843 ATGCAATCTTCTGATATATACCATCAACGAGAAAGTATATAGTTTACAACACAGCTAAATGGC 902  
Db |||||||  
Qy 721 ATGCAATCTTCTGATATATACCATCAACGAGAAAGTATATAGTTTACAACACAGCTAAATGGC 780  
Db |||||||  
Qy 903 TACACCAATTAACCTACCATAACGGTACACCTACTCAAAATAATACCAAAAGGCTAGACCT 962  
Db |||||||  
Qy 781 TACACCAATTAACCTACCATAACGGTACACCTACTCAAAATAATACCAAAAGGCTAGACCT 840  
Db |||||||  
Qy 963 CTGCTCACTAGGCCCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAATA 1022  
Db |||||||  
Qy 841 CTGCTCACTAGGCCCTAGCTTTTACATCAAAAGCAATGACTATAGAAAGTAAAAATA 900  
Db |||||||  
Qy 1023 CTAGAAAAGCAATGATTGCGCTCATTCG 1051  
Db |||||||  
Qy 901 CTAGAAAAGCAATGATTGCGCTCATTCG 929  
Db |||||||

## RESULT 9

US-08-482-785-14  
Sequence 14, Application US/08482785  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,785  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE: 24-FEB-1995  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 937 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..819  
US-08-482-785-14

Query Match 78.8%; Score 853.8; DB 9; Length 937;  
Best Local Similarity 95.7%; Pred. No. 2.3e-186;  
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTTCTAAAAAATGTGCGCTAGTAAAA 182  
Db |||||||  
Qy 183 TTTTCAATGTAGTCTTTGTATCAGCCACAATGGCTGTAAACAACAGTCACACTTTGAAAT 242  
Db |||||||  
Qy 61 TTCTCCATGTTGCTCTGTTTCCGTCACCATGGCTTACCACCGTTACCTCGGAAAC 120  
Db |||||||  
Qy 243 ACTGCATCTGCACGACGACAAACACAGGCTCTCAATGATGTTGTTCTAAATGATGGCGCAGC 302  
Db |||||||  
Qy 121 ACCGCTCTGGC---TCAGACACAGGCTCTCAATGATGTTGTTCTAAATGATGGCGCAGC 177  
Db |||||||  
Qy 303 AAGTACCTTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCTAACTATTACAAACT 362  
Db |||||||  
Qy 178 AAGTACCTTAAACGAAGCATTAGCTTGGACATTCATGACAGTCTCTAACTATTACAAACT 237  
Db |||||||  
Qy 363 TTAGGTACTAGTCAGATTAATCTCCAGCACTCTTTCTAAAGCAGGAGATTTCTCTATAGC 422  
Db |||||||  
Qy 238 TTAGGTACTAGTCAGATTAATCTCCAGCACTCTTTCTAAAGCAGGAGATTTCTCTATAGC 297  
Db |||||||  
Qy 423 AAATTAGATGATTAGGAGGACGCGTACTGCTAGAGGTACATTTGACTTTATGCCAATGTT 482  
Db |||||||  
Qy 298 AAATTAGATGATTAGGAGGACGCGTACTGCTAGAGGTACATTTGACTTTATGCCAATGTT 357  
Db |||||||  
Qy 483 GAAGGTAGCTACGGTGTGTAGACAATCTTTCGGTAAATAATCAAAACCCCGCAGGATGGACT 542  
Db |||||||  
Qy 358 GAAGGTAGCTACGGTGTGTAGACAATCTTTCGGTAAATAATCAAAACCCCGCAGGATGGACT 417  
Db |||||||  
Qy 543 GGAACCCCTAATCATGTCAAAATATAAAATGGTTAAATGCTCTATCTTATGTCGGA 602  
Db |||||||  
Qy 418 GGAACCCCTAATCATGTCAAAATATAAAATGGTTAAATGCTCTATCTTATGTCGGA 477  
Db |||||||  
Qy 603 GATTTCGGAATAGAAGTCATCTTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 662  
Db |||||||  
Qy 478 GATTTCGGAATAGAAGTCATCTTCATTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 537  
Db |||||||  
Qy 663 AATCCGTTACAGAAACAGTCACCAAAATGTAGAGTCTGTGACCAAAAGGCGGCATG 722  
Db |||||||  
Qy 538 AATCCGTTACAGAAACAGTCACCAAAATGTAGAGTCTGTGACCAAAAGGCGGCATG 597  
Db |||||||  
Qy 723 CGCTATACCGAAACAAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 782  
Db |||||||  
Qy 598 CGCTATACCGAAACAAAGAGCTCAAGAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 657  
Db |||||||  
Qy 783 TATGAAGTCGCTCCAAATCTACAACGACAGAGTGTGATTCGAAGAGCTGTGCTGGTATCA 842  
Db |||||||

Db 658 TATGAAGTCCTCCAACTTCTAACACGACGAGCTGATTCACAGAGCTGCTGGTATCA 717  
Qy 843 ATGCAATCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGCG 902  
Db 718 ATGCAATCTTCTGATATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGCG 777  
Qy 903 TACACCATTAATACCATTAACCGGTACACCTTCTCAAAAAATAATACCAAAAGGCTAGACCT 962  
Db 778 TACACCATTAATACCATTAACCGGTACACCTTCTCAAAAAATAATACCAAAAGGCTAGACCT 837  
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 1022  
Db 838 CTGCTCACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 897  
Qy 1023 CTAGAAAAAGCAATGATTCGCGTCATTGC 1051  
Db 898 CTAGAAAAAGCAATGATTCGCGTCATTGC 926

## RESULT 10

US-09-119-900-14  
; Sequence 14, Application US/09119900  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 937 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..819  
US-09-119-900-14

Query Match 78.8%; Score 853.8; DB 16; Length 937;  
Best Local Similarity 95.7%; Pred. No. 2.3e-186;

Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;  
Qy 123 AAGCATATGATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATCTCGGCTAGTAAAA 182  
Db 1 ATGATCGAAGACCTGCTGGGTTCGGTCTGTTTCTCCAAAAAATCCCGCTCTGGTTAAA 60  
Qy 183 TTTTCAATGGTGTAGTCTTGTATCAGCCACCAATGGCTGTAAACACAGTGCACACTTGAAT 242  
Db 61 TTCTTCAATGGTGTCTTGTGTTTCGGTACCATGGCTGTATCCACCGGTACCCCTGGA 120  
Qy 243 ACTGCTACGACGACAAACACAGGTCTCAAAATGATGTTCTTAAATGATGGCGAAGC 302  
Db 121 ACCGCTCTGGC---TCAGACACAGGTCTCAATGATGTTCTTAAATGATGGCGAAGC 177  
Qy 303 AAGTACCTAAACGAAGCAATTAGCTTGGACATTCATCAATGACAGTCTTAACTAATTA 362  
Db 178 AAGTACCTAAACGAAGCAATTAGCTTGGACATTCATCAATGACAGTCTTAACTAATTA 237  
Qy 363 TTAGGTACTAGTCTCAGATTACTCCAGCACTCTTTTCTTAAAGCAGGAGATATCTCTAT 422  
Db 238 TTAGGTACTAGTCTCAGATTACTCCAGCACTCTTTTCTTAAAGCAGGAGATATCTCTAT 297  
Qy 423 AAATTAGATGATTTAGGAAGGACGCGTACTCTGTAGAGGTACATTTGACTTTATGCAAT 482  
Db 298 AAATTAGATGATTTAGGAAGGACGCGTACTCTGTAGAGGTACATTTGACTTTATGCAAT 357  
Qy 483 GAAGGTAGCTACGGTGTAGACAAATCTTTCGGTAAAAAATCAAAACCCCGCAGGATGG 542  
Db 358 GAAGGTAGCTACGGTGTAGACAAATCTTTCGGTAAAAAATCAAAACCCCGCAGGATGG 417  
Qy 543 GGAACCCCTAAATCATGTCAAAATATAAAATTTGAATGTTTAAATGTTTAAATGGTCT 602  
Db 418 GGAACCCCTAAATCATGTCAAAATATAAAATTTGAATGTTTAAATGGTCTTCTTATG 477  
Qy 603 GATTTCTGGAATAGAAAGTCAATCTCATTTGAGATAGTCTCGGTGGAGATGCACCTCAG 662  
Db 478 GATTTCTGGAATAGAAAGTCAATCTCATTTGAGATAGTCTCGGTGGAGATGCACCTCAG 537  
Qy 663 AATGCGGTTACAGGAACAGTACCCAAATGTAGGAGGTCTGTGACCAAAAGGCGCGCAT 722  
Db 538 AATGCGGTTACAGGAACAGTACCCAAATGTAGGAGGTCTGTGACCAAAAGGCGCGCAT 597  
Qy 723 CGCTATACCGAACAAGAGCTCAAGAAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 782  
Db 598 CGCTATACCGAACAAGAGCTCAAGAAATGGTTAGAGCAAAATCGTGATGGCTATCTTTAT 657  
Qy 783 TATGAAGTCTGCTCCAAATCTAACCGACGAGTGTGATTCAGAGCTGTCTGGTATCA 842  
Db 658 TATGAAGTCTGCTCCAAATCTAACCGACGAGTGTGATTCAGAGCTGTCTGGTATCA 717  
Qy 843 ATGCAATCTTCTGATTAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 902  
Db 718 ATGCAATCTTCTGATTAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATG 777  
Qy 903 TACACCATTAATACCATTAACCGGTACACCTTCTCAAAAAATAATACCAAAAGGCTAG 962  
Db 778 TACACCATTAATACCATTAACCGGTACACCTTCTCAAAAAATAATACCAAAAGGCTAG 837  
Qy 963 CTGCTCACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 1022  
Db 838 CTGCTCACTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 897  
Qy 1023 CTAGAAAAAGCAATGATTCGCGTCATTGC 1051  
Db 898 CTAGAAAAAGCAATGATTCGCGTCATTGC 926

## RESULT 11

US-10-453-032-14  
; Sequence 14, Application US/10453032  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.



```

; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/453,032
; FILING DATE: 03-JUNE-2003
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE: 23-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 937 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..819
; US-10-453-032-14

Query Match 78.8%; Score 853.8; DB 53; Length 937;
Best Local Similarity 95.7%; Pred.No. 2.3e-186;
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;

Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAAAAATGTCGGCTAGTAAA 182
Db 1 ATGGATCCGAACCTGCTGGGTCCCGTCGTGTTTCTCCAAAAAATGCCGCTGTTAA 60
Qy 183 TTTTCAATGTAGTCTTTGATCAGCCACAAATGGCTGTAAACACAGTCACACTTGAATA 242
Db 61 TTCTCCATGTTGCTCTGTTTCGGCTTACCATGGCTGTAAACACCGTTACCCCTGGAAC 120
Qy 243 ACTGCACTGGCAGCAGCAAAACACAGGTCCTAAATGATGTTGTTCTAAATGATGCGCAAG 302
Db 121 ACCGCTCTGGC---TCAGACACAGGTCCTAAATGATGTTGTTCTAAATGATGCGCAAG 177
Qy 303 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGTCCTTAACATTAACAAAAC 362
Db 178 AAGTACCTAAACGAAGCATTTAGCTTGGACATTCATGACAGTCCTTAACATTAACAAAAC 237
Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATTTCTCTATAGC 422
Db 238 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATTTCTCTATAGC 297
Qy 423 AAATTAGATGATGATGAGGAGCGCTACTGCTAGAGGTACATTTGACTTATGCCAATGTT 482
Db 298 AAATTAGATGATGATGAGGAGCGCTACTGCTAGAGGTACATTTGACTTATGCCAATGTT 357

483 GAAGGTAGTACGGTGTGTAGACAATCTTTCCGGTAAAAATCAAAACCCCGCAGGATGACT 542
358 GAAGGTAGTACGGTGTGTAGACAATCTTTCCGGTAAAAATCAAAACCCCGCAGGATGACT 417
543 GGAACCCCTAATCATGTCAAAATATAAAATTTGAATGTTAAATGTTCTTATGTCGGA 602
418 GGAACCCCTAATCATGTCAAAATATAAAATTTGAATGTTAAATGTTCTTATGTCGGA 477
603 GATTTCCTGGAATAGAGTCATCTCATTCGTCAGATAGTCTCCGTCGAGATGCACTCAGAGTC 662
478 GATTTCCTGGAATAGAGTCATCTCATTCGTCAGATAGTCTCCGTCGAGATGCACTCAGAGTC 537
663 AATGCGGTTTACAGGAACACGTACCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCATG 722
538 AATGCGGTTTACAGGAACACGTACCCAAAATGTAGGAGGTGCTGACCAAAAAGCGCGCATG 597
723 CGCTATATACCGAACAAAAGAGCTCAAGAAATGTTTGAAGCAAAATCGTGATGGCTATCTTTAT 782
598 CGCTATATACCGAACAAAAGAGCTCAAGAAATGTTTGAAGCAAAATCGTGATGGCTATCTTTAT 657
783 TATGAAGTCGCTCCAAATCTACAAACGAGAGAGTTGATTCCAAAGAGCTGTGCGGTATCA 842
658 TATGAAGTCGCTCCAAATCTACAAACGAGAGAGTTGATTCCAAAGAGCTGTGCGGTATCA 717
843 ATGCAATCTTCTGATTAATACCATCAACGAGAGAAAGTATTAGTTTACAAACACAGCTAATGGC 902
718 ATGCAATCTTCTGATTAATACCATCAACGAGAGAAAGTATTAGTTTACAAACACAGCTAATGGC 777
903 TACACCAATTAACCTAACCGTACACCTACTCAAAAATTAATACCAAAAGGCTAGACCT 962
778 TACACCAATTAACCTAACCGTACACCTACTCAAAAATTAATACCAAAAGGCTAGACCT 837
963 CTGCTCACTAGGCTAGCTATTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 1022
838 CTGCTCACTAGGCTAGCTATTTTACATCAAAAAGCAATGACTATAGAAAAGTAAAAATA 897
1023 CTAGAAAAGCAATGATTCGGTCATTGC 1051
898 CTAGAAAAGCAATGATTCGGTCATTGC 926

RESULT 12
US-08-482-785-10
; Sequence 10, Application US/08482785
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,785
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Michael B.
REGISTRATION NUMBER: 32,612
REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
US-08-482-785-10

Query Match      18.5%; Score 200; DB 9; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAGCGCTCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAAC 60
DB 1 GACAAGCGCTCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAAC 60
QY 61 TATTGATAAACTAGTTAAGTAAAGCGTATATCTAGTGGTAAATTTAGAAAAGAGG 120
DB 61 TATTGATAAACTAGTTAAGTAAAGCGTATATCTAGTGGTAAATTTAGAAAAGAGG 120
QY 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
DB 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
QY 181 AATTTTCAATGGTAGCTCTT 200
DB 181 AATTTTCAATGGTAGCTCTT 200
```

```
RESULT 13
US-09-119-900-10
Sequence 10, Application US/09119900
GENERAL INFORMATION:
APPLICANT: Adams, Craig W.
APPLICANT: Pang, Patty P.-Y.
APPLICANT: Belei, Marina
TITLE OF INVENTION: Recombinant DNase B Derived from
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheldon & Mak
STREET: 225 South Lake Avenue, Ninth Floor
CITY: Pasadena
STATE: California
COUNTRY: USA
ZIP: 91001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/119,900
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/082,845
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Michael B.
REGISTRATION NUMBER: 32,612
```

```
REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
TELEFAX: (818) 795-6321
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
US-09-119-900-10

Query Match      18.5%; Score 200; DB 16; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAGCGCTCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAAC 60
DB 1 GACAAGCGCTCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAAC 60
QY 61 TATTGATAAACTAGTTAAGTAAAGCGTATATCTAGTGGTAAATTTAGAAAAGAGG 120
DB 61 TATTGATAAACTAGTTAAGTAAAGCGTATATCTAGTGGTAAATTTAGAAAAGAGG 120
QY 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
DB 121 ACAAGCATATGAATCTACTTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
QY 181 AATTTTCAATGGTAGCTCTT 200
DB 181 AATTTTCAATGGTAGCTCTT 200

RESULT 14
US-10-453-032-10
Sequence 10, Application US/10453032
GENERAL INFORMATION:
APPLICANT: Adams, Craig W.
APPLICANT: Pang, Patty P.-Y.
APPLICANT: Belei, Marina
TITLE OF INVENTION: Recombinant DNase B Derived from
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheldon & Mak
STREET: 225 South Lake Avenue, Ninth Floor
CITY: Pasadena
STATE: California
COUNTRY: USA
ZIP: 91001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/453,032
FILING DATE: 03-JUNE-2003
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/082,845
FILING DATE: 23-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Michael B.
REGISTRATION NUMBER: 32,612
REFERENCE/DOCKET NUMBER: 9521
TELECOMMUNICATION INFORMATION:
TELEPHONE: (818) 796-4000
```

```
TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
US-10-453-032-10

Query Match      18.5%; Score 200; DB 53; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e-35;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACACGCCCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAC 60
   |||||
Db 1 GACACGCCCTCTTTTCTCCTTACTATCTCTTAAATTTTCATATTTTAAAAAAC 60
   |||||

QY 61 TATTGATAAACTAGTTAAGTAAGGCTATCTATGTTAGTTAGCGAAATTAGAAAAGAGG 120
   |||||
Db 61 TATTGATAAACTAGTTAAGTAAGGCTATCTATGTTAGTTAGCGAAATTAGAAAAGAGG 120
   |||||

QY 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
   |||||
Db 121 ACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180
   |||||

QY 181 AATTTTCAATGGTAGCTCTT 200
   |||||
Db 181 AATTTTCAATGGTAGCTCTT 200
   |||||

RESULT 15
US-08-482-785-12
; Sequence 12, Application US/08482785
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,785
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Synthetic primer
US-08-482-785-12

Query Match      9.8%; Score 106.6; DB 9; Length 182;
Best Local Similarity 77.3%; Pred. No. 4.3e-14;
Matches 143; Conservative 0; Mismatches 39; Indels 3; Gaps 1;

QY 118 AGGACACGACATATGTAATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAG 177
   |||||
Db 1 AGGCAATGGATCCGAACCTGCTGGGTTCCGCTGCTGTTTCTCAAAAAAATGCGGTCTGG 60
   |||||

QY 178 TAAAAATTTTCAATGGTAGCTCTTTGTATCAGCCCAATGGCTGTAACACAGTCACACTTG 237
   |||||
Db 61 TTAATTTCTCCATGGTTGCTCTGTTTCCGCTACCATGGCTGTACCACGTTACCCCTGG 120
   |||||

QY 238 AAAATACTGCACTGGGACGACAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCG 297
   |||||
Db 121 AAAACACCGCTCTGGC---TCAGACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCG 177
   |||||

QY 298 CAAGC 302
   |||||
Db 178 CAAGC 182
   |||||

Search completed: January 5, 2004, 16:40:20
Job time : 3248 secs
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: January 5, 2004, 13:55:47 ; Search time 3023 Seconds  
(without alignments)  
863.142 Million cell updates/sec

Title: US-08-482-785-7  
Perfect score: 1083  
Sequence: 1 GACACGCCCTCTTTTCT.....TGTCAAAAGCAAAAGC 1083

## Scoring table:

IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4485014 seqs, 1204653517 residues

Total number of hits satisfying chosen parameters: 8970028

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA New:  
1: /cgn2\_6/ptodata/2/pna/PCT\_NEW\_COMB.seq:  
2: /cgn2\_6/ptodata/2/pna/US06\_NEW\_COMB.seq:  
3: /cgn2\_6/ptodata/2/pna/US07\_NEW\_COMB.seq:  
4: /cgn2\_6/ptodata/2/pna/US08\_NEW\_COMB.seq:  
5: /cgn2\_6/ptodata/2/pna/US09\_NEW\_COMB.seq:  
6: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq:  
7: /cgn2\_6/ptodata/2/pna/US10\_NEW\_COMB.seq2:  
8: /cgn2\_6/ptodata/2/pna/US60\_NEW\_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	4	US-08-472-630-7
2	922.6	85.2	940	4	US-08-472-630-11
3	853.8	78.8	937	4	US-08-472-630-14
4	200	18.5	200	4	US-08-472-630-10
5	106.6	9.8	182	4	US-08-472-630-12
6	46.2	4.3	6106	7	US-10-221-714A-151
7	42.4	3.9	7503	7	US-10-433-793-39
8	42.4	3.9	167163	7	US-10-394-948-31
9	41.4	3.8	5413	7	US-10-221-714A-417
10	41.2	3.8	37184	7	US-10-433-793-108
11	40.8	3.8	8238	7	US-10-240-454A-43
12	40.2	3.7	832	6	US-10-664-025-2813
13	40	3.7	3814	7	US-10-473-126-150
14	40	3.7	133801	8	US-08-487-610-19911
15	39.4	3.6	7057	7	US-10-221-714A-324
16	39.2	3.6	6192	7	US-10-221-714A-408
17	39.2	3.6	10609	7	US-10-221-714A-81
18	39.2	3.6	15592	7	US-10-221-714A-175
19	39	3.6	6179	7	US-10-221-714A-66
20	39	3.6	14362	7	US-10-433-793-77
21	39	3.6	14708	7	US-10-221-714A-499
22	39	3.6	14708	7	US-10-240-589C-133
23	39	3.6	83391	7	US-10-433-793-123
24	38.8	3.6	6040	7	US-10-240-589C-110
25	38.8	3.6	6676	7	US-10-433-793-43

C 26 38.8 3.6 13131 7 US-10-240-589C-58 Sequence 58, Appl  
C 27 38.8 3.6 544420 8 US-60-487-610-19233 Sequence 19233, A  
C 28 38.8 3.6 545735 8 US-60-500-337-19015 Sequence 19015, A  
C 29 38.4 3.5 3814 7 US-10-473-126-296 Sequence 296, App  
C 30 38.4 3.5 56153 7 US-10-221-714A-520 Sequence 520, App  
C 31 38.4 3.5 8059021 5 US-09-947-914-53 Sequence 53, Appl  
C 32 38.2 3.5 6470 7 US-10-433-793-121 Sequence 121, Appl  
C 33 38.2 3.5 8753 7 US-10-240-589C-80 Sequence 80, Appl  
C 34 38.2 3.5 193757 7 US-10-719-993-6939 Sequence 6939, App  
C 35 38.2 3.5 193757 8 US-60-519-270-2355 Sequence 2355, App  
C 36 38.2 3.5 1790242 7 US-10-719-993-6940 Sequence 6940, App  
C 37 38.2 3.5 1790242 8 US-60-519-270-2356 Sequence 2356, App  
C 38 38 3.5 201 8 US-60-487-610-82379 Sequence 82379, A  
C 39 38 3.5 12269 7 US-10-221-714A-194 Sequence 194, App  
C 40 38 3.5 40324 7 US-10-433-793-180 Sequence 180, App  
C 41 38 3.5 200335 8 US-60-500-337-19086 Sequence 19086, A  
C 42 38 3.5 1790242 7 US-10-719-993-6940 Sequence 6940, App  
C 43 38 3.5 1790242 8 US-60-519-270-2356 Sequence 2356, App  
C 44 38 3.5 8059021 5 US-09-947-914-53 Sequence 53, Appl  
C 45 37.8 3.5 6219 7 US-10-240-454A-19 Sequence 19, Appl

## ALIGNMENTS

RESULT 1  
US-08-472-630-7  
; Sequence 7, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1083 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:



Db 61 TTTTCAATGGTAGCTCTTGATCAGCCACCAATGGCTGAACAACAGTCACACTTGAAAT 120  
Qy 243 ACTGCACTGGCAGCAGCAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGCAAGC 302  
Db 121 ACTGCACTGGCAGCAGCAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGCAAGC 180  
Qy 303 AAGTACCTAAACGAAGCATTAGCTTGGACATTCGAATGACAGTCTCTAACTATTACAAACT 362  
Db 181 AAGTACCTAAACGAAGCATTAGCTTGGACATTCGAATGACAGTCTCTAACTATTACAAACT 240  
Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTATAGC 422  
Db 241 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTATAGC 300  
Qy 423 AAATTAGATGATTAGGAAGACCGTACTGCTAGAGGTACATTGACTTATGCAATGTT 482  
Db 301 AAATTAGATGATTAGGAAGACCGTACTGCTAGAGGTACATTGACTTATGCAATGTT 360  
Qy 483 GAAGGTAGTACGGTGTGACAAATTAATAATGTTAAATGGTTAAATGGTCTATCTTATGCGGA 542  
Db 361 GAAGGTAGTACGGTGTGACAAATTAATAATGTTAAATGGTTAAATGGTCTATCTTATGCGGA 420  
Qy 543 GGAACCTTAATCATGTCAATATAATAATGTTAAATGGTTAAATGGTCTATCTTATGCGGA 602  
Db 421 GGAACCTTAATCATGTCAATATAATAATGTTAAATGGTTAAATGGTCTATCTTATGCGGA 480  
Qy 603 GATTCTGGAATAGAGTCAATCTCATGTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 662  
Db 481 GATTCTGGAATAGAGTCAATCTCATGTCAGATAGTCTCGGTGGAGATGCACTCAGAGTC 540  
Qy 663 AATGCCGTACAGGACACATCCCAAAATGTAGGAGTCTGACCAAAAGGCGGCATG 722  
Db 541 AATGCCGTACAGGACACATCCCAAAATGTAGGAGTCTGACCAAAAGGCGGCATG 600  
Qy 723 CGCTATACCGAAACAGAGCTCAAGATGTTAGAACAAATCGTGATGGCTATCTTTAT 782  
Db 601 CGCTATACCGAAACAGAGCTCAAGATGTTAGAACAAATCGTGATGGCTATCTTTAT 660  
Qy 783 TATGAAGTCGCTCAATCTCAACGACGAGTGTATCCAAAGAGCTGTCGTGGTATCA 842  
Db 661 TATGAAGTCGCTCAATCTCAACGACGAGTGTATCCAAAGAGCTGTCGTGGTATCA 720  
Qy 843 ATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGGC 902  
Db 721 ATGCAATCTCTGATAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTAATGGC 780  
Qy 903 TACACCAATTAACCTACCATACGAGAAAGTATTAGTTTACACACAGCTAATGGC 962  
Db 781 TACACCAATTAACCTACCATACGAGAAAGTATTAGTTTACACACAGCTAATGGC 840  
Qy 963 CTGCTCAGTGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAATGAAAAATA 1022  
Db 841 CTGCTCAGTGGCTAGCTTTTACATCAAAAAGCAATGACTATAGAAAATGAAAAATA 900  
Qy 1023 CTAGAAAAGCAATGATGCGCTATTGC 1051  
Db 901 CTAGAAAAGCAATGATGCGCTATTGC 929

## RESULT 3

US-08-472-630-14

; Sequence 14, Application US/08472630

; GENERAL INFORMATION:

; APPLICANT: Adams, Craig W.

; Belci, Marina

; Pang, Patty P.-Y.

; TITLE OF INVENTION: Recombinant DNase B Derived from

; Streptococcus pyogenes

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheldon &amp; Mak

; STREET: 225 South Lake Avenue, Ninth Floor

; CITY: Pasadena

; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472.630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082.845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 937 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 1..819  
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
; US-08-472-630-14

Query Match 78.8%; Score 853.8; DB 4; Length 937;  
Best Local Similarity 95.7%; Pred. No. 1.2e-246;  
Matches 889; Conservative 0; Mismatches 37; Indels 3; Gaps 1;  
Qy 123 AAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAATAATGTCGGCTAGTAAA 182  
Db 1 ATGGATCCGAACCTGCTGGTTCCTCCAAAAAATGCGCTCTGGTTAA 60  
Qy 183 TTTTCAATGGTAGCTCTTGATCAGCCACAAATGGCTGTAAACAGTCACACTTGAAT 242  
Db 61 TTCTCCATGGTGTGCTCTGGTTTCCGCTACCATGGCTGTACCACCGTTACCTGGAAC 120  
Qy 243 ACTGCACTGGCAGCAGCAAAACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGAAGC 302  
Db 121 ACCGCTCTGGC---TCAGACACAGGTCTCAAAATGATGTTGTTCTAAATGATGGCGAAGC 177  
Qy 303 AAGTACCTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGTCTTAACTATTACAAACT 362  
Db 178 AAGTACCTAAACGAAGCATTAGCTTGGACATTCATCAATGACAGTCTTAACTATTACAAACT 237  
Qy 363 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTATAGC 422  
Db 238 TTAGGTACTAGTCAGATTACTCCAGCACTCTTTCTTAAAGCAGGAGATATTTCTATAGC 297  
Qy 423 AAATTAGATGATTAGGAAGACCGTACTGCTAGAGGTACATTGACTTATGCAATGTT 482  
Db 298 AAATTAGATGATTAGGAAGACCGTACTGCTAGAGGTACATTGACTTATGCAATGTT 357  
Qy 483 GAAGGTAGCTACGGTGTGACAAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 542  
Db 358 GAAGGTAGCTACGGTGTGACAAATCTTTTCGGTAAAAATCAAAACCCCGCAGGATGACT 417  
Qy 543 GGAACCTTAATCATGTCAATATAATAATGTTAAATGGTTAAATGGTCTATCTTATGCGGA 602

Db 418 GGAACCCCTAATCATGTCAAAATATAAAATGAATGGTTAAATGGTCTCTCTATGTCGGA 477  
Qy 603 GATTTCGGAATAGAGTCATCTCATTCGAGATAGTCTCGTGAGATGACCTCAAGTTC 662  
Db 478 GATTTCGGAATAGAGTCATCTCATTCGAGATAGTCTCGTGAGATGACCTCAAGTTC 537  
Qy 663 AATGCCGTTTACAGGAACACGCTACCCAAATGTAGGAGGTCGTGACCAAAAAGCGGCATG 722  
Db 538 AATGCCGTTTACAGGAACACGCTACCCAAATGTAGGAGGTCGTGACCAAAAAGCGGCATG 597  
Qy 723 CGGTATACCAACAAAGAGCTCAAGATGTTTGAAGCAAAATCGTGATGCTATCTTTAT 782  
Db 598 CGGTATACCAACAAAGAGCTCAAGATGTTTGAAGCAAAATCGTGATGCTATCTTTAT 657  
Qy 783 TATGAAGTCGCTCCCAATCTACACGCGAGAGGTCGATTCACAGAGCTGCTGCTATCA 842  
Db 658 TATGAAGTCGCTCCCAATCTACACGCGAGAGGTCGATTCACAGAGCTGCTGCTATCA 717  
Qy 843 ATGCAATCTCTGATATACCAATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGCG 902  
Db 718 ATGCAATCTCTGATATACCAATCAACGAGAAAGTATTAGTTTACAAACAGCTAATGCG 777  
Qy 903 TACACCAATTAACCTACCAATCAACGCTACCTCAACCAATTAACCAACAGCTAATGCG 962  
Db 778 TACACCAATTAACCTACCAATCAACGCTACCTCAACCAATTAACCAACAGCTAATGCG 837  
Qy 963 CTGCTCAGTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 1022  
Db 838 CTGCTCAGTAGGCTAGCTTTTACATCAAAAAAGCAATGACTATAGAAAGTAAAAATA 897  
Qy 1023 CTAGAAAAAGCAATGATGCGCTGCTATGC 1051  
Db 898 CTAGAAAAAGCAATGATGCGCTGCTATGC 926

RESULT 4

US-08-472-630-10  
; Sequence 10, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 796-6321  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 200 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORGANISM: Streptococcus pyogenes  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-472-630-10

Query Match 18.5%; Score 200; DB 4; Length 200;  
Best Local Similarity 100.0%; Pred. No. 4.9e-50;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GACAAGCCCTCTCTTTTCTCCCTACTATCTCTTAATTTTCATATTTTAAAAAAC 60  
Db 1 GACAAGCCCTCTCTTTTCTCCCTACTATCTCTTAATTTTCATATTTTAAAAAAC 60  
Qy 61 TATTGATAAACTAGTTAAAGTAAGCGTATATCTAGTTCGCGAAATAGAAAAAGG 120  
Db 61 TATTGATAAACTAGTTAAAGTAAGCGTATATCTAGTTCGCGAAATAGAAAAAGG 120  
Qy 121 ACAAGCATATGATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180  
Db 121 ACAAGCATATGATCTACTTGGATCAAGACGGGTTTTTCTAAAAAATGTCGGCTAGTAA 180  
Qy 181 AATTTTCAATGGTAGCTCTT 200  
Db 181 AATTTTCAATGGTAGCTCTT 200

RESULT 5

US-08-472-630-12  
; Sequence 12, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 796-6321  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 182 base pairs  
; TYPE: nucleic acid



```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Synthetic primer
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-08-472-630-12

Query Match          9.8%; Score 106.6; DB 4; Length 182;
Best Local Similarity 77.3%; Pred. No. 6.2e-22;
Matches 143; Conservative 0; Mismatches 39; Indels 3; Gaps 1;

Qy 118 AGGACAAGCATATGAATCTACTTGGATCAAGACGGGTTTTTCTTAAAAAATGTCGGCTAG 177
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 178 TAAATTTTCAATGGTAGCTCTTGATCAGCCACAATGGCTGTAAACAACAGTCACTTG 237
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 61  TTAATTTCTCCATGGTTGCTCTGTTTCCGCTACCATGGCTGTACCACGGTTACCCCTGG 120
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 238 AAAATCTGCACCTGGCAGCAACACACAGGCTCAAAATGATGTTCTTAAATGATGGCG 297
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 121 AAACACCGCTCTGGC---TCAGACAGGTTCTAAATGATGTTGTTCTTAAATGATGGCG 177
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 298 CAAGC 302
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 178 CAAGC 182
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 6
US-10-221-714A-151/c
; Sequence 151, Application US/10221714A
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: tumor suppressor genes and oncogenes
; FILE REFERENCE: 5013.1005
; CURRENT APPLICATION NUMBER: US/10/221.714A
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: PCT/EP01/02955
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: DE 10013847.0
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 540
; SEQ ID NO 151
; LENGTH: 6106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-714A-151

Query Match          4.3%; Score 46.2; DB 7; Length 6106;
Best Local Similarity 48.3%; Pred. No. 0.0041;
Matches 158; Conservative 0; Mismatches 168; Indels 1; Gaps 1;

Qy 36  TAATTTTTCATATTTTAAAAAACTATTCATAAACTAGTTAAGTAAGCGTATACTATCG 95
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 5349 TAACTTTAAACTCTTAAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5290
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 96  TTAGTTAGCGAAATTAGAAAAGAGGACAAGCATATGAATCTACTTGGATCAAGACGGGTT 155
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

NAME/KEY: misc\_feature  
LOCATION: (1)...(167163)  
OTHER INFORMATION: n = A,T,C or G  
US-10-394-948-31

Query Match 3.8%; Score 42.4; DB 7; Length 167163;  
Best Local Similarity 50.0%; Pred. No. 0.22;  
Matches 106; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY 11 TCTTTTCTCTTACTATCTCTTAAATTTTCATATTTTAAATAAATTCATATAA 70  
DB 137416 TCAATTAATCAATCTATCCAAAGTACTAAATTTTCATAAAACAATCTTTATAAACT 137357  
QY 71 CTAGTTAAGTAAGCGGTATATCTATGTTAGTCCGAAATAGAAAAGAGCAACATAT 130  
DB 137356 CAATTTTCATATTGTTATGTCGTACTTAATTAACCAACACAACTTTTGGCAAGCAT 137297  
QY 131 GAATCTACTTGATCAAGAGCGGTTTTTCTAAATAATGTGGCTAGTAAATTTTCAAT 190  
DB 137296 AATCAACTAGTAGTAACAAAGCTTCTGGGCACCAATATCAATAGCTAACATTTATTG 137237  
QY 191 GGTAGCTCTTGATCAGCCACAAATGGCTGTAA 222  
DB 137236 AGTGTTATCGTATCATAAACACGTCTTAA 137205

RESULT 9  
US-10-221-714A-417/c  
Sequence 417, Application US/10221714A  
GENERAL INFORMATION:

APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
TITLE OF INVENTION: Diagnosis of Diseases Associated with  
FILE OF INVENTION: tumor suppressor genes and oncogenes  
FILE REFERENCE: 5013.1005  
CURRENT APPLICATION NUMBER: US/10/221,714A  
CURRENT FILING DATE: 2003-01-21

PRIOR APPLICATION NUMBER: PCT/EP01/02955  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: DE 10013847.0  
PRIOR FILING DATE: 2000-03-15  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 540

SEQ ID NO 417  
LENGTH: 5413  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-221-714A-417

Query Match 3.8%; Score 41.4; DB 7; Length 5413;  
Best Local Similarity 52.6%; Pred. No. 0.11;  
Matches 90; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 911 TAATCACTAACGGTACCTACTCAAAATAATACCAAAAGCTAGACCTCTGCTCAC 970  
DB 1096 TAAATAACAAACGAAACTCCATCTCAAAACAAACTTAAACACGATCACTACTTA 1037  
QY 971 TAGCCCTAGCTTTTACATCAAAAGCATGCTATAGAAAGTAAATCTAGAAA 1030  
DB 1036 TAATCCACACCTTTAAACCAAAATAATAATCACTAAATCAAAAATTCAAAC 977  
QY 1031 AGCAATGATTCGGCTCATTCCTTTTATGATATTTGTGCAAAAGCAAAA 1081

Db 976 AACCACTTAACCAACAATAATAAAACCCTATCTCTACTAAATAACAAA 926

## RESULT 10

US-10-433-793-108/c  
Sequence 108, Application US/10433793  
GENERAL INFORMATION:  
APPLICANT: Epigenomics AG  
TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/433,793  
CURRENT FILING DATE: 2003-06-06  
NUMBER OF SEQ ID NOS: 212  
SEQ ID NO 108  
LENGTH: 37184  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-433-793-108

Query Match 3.8%; Score 41.2; DB 7; Length 37184;  
Best Local Similarity 50.5%; Pred. No. 0.28;  
Matches 100; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 838 TATCAATGCAATCTCTGTATAATACCATCAACGAGAAAGTATTAGTTTACACACAGCTA 897  
DB 29539 TCTATTCTTAATCCCATATATTAACCAAAACAAATTAATCCCACTATATCCCACTATACTAATA 29480  
QY 898 ATGGCTACACCAATTAATACCTAACGCTACACTCTCAAAATAATATACCAAAAGCTA 957  
DB 29479 ATCAAAAACCAATTTCCCAAAAATAATTAATTAATCTTATCCAAATTAATAGCAAAATAA 29420  
QY 958 GACCTCTGCTACTAGGCTAGCTTTTATACATCAAAAGCAATGACATATAGAAAGTAA 1017  
DB 29419 TCTCAAAACTAAATTAATCTTACCCCACTAATTTTAACTACTAATAATAATAAAAAA 29360  
QY 1018 AATATCTAGAAAAAGCAA 1035  
DB 29359 ATAAAAAAAACCTCCCA 29342

## RESULT 11

US-10-240-454A-43/c  
Sequence 43, Application US/10240454A  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
APPLICANT: BERLIN, Kurt  
TITLE OF INVENTION: Diagnosis of Diseases Associated with Metabolism  
FILE REFERENCE: 5013.1010  
CURRENT APPLICATION NUMBER: US/10/240,454A  
CURRENT FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: PCT/EP01/04016  
PRIOR FILING DATE: 2001-04-06  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 68  
SEQ ID NO 43  
LENGTH: 8238  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-240-454A-43

Query Match 3.8%; Score 40.8; DB 7; Length 8238;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:38:21 ; Search time 323.403 Seconds  
(without alignments)  
762.478 Million cell updates/sec

Title: US-08-482-785-8

Perfect score: 1418

Sequence: 1 MNLGSRVFSKRLVKFS.....VYNTANGTYNYHNGPTQK 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.\*

```

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/2/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/2/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/2/paa/US60_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1418	100.0	271	8	US-08-482-785-8
					Sequence 8, Appli

2	1418	100.0	271	15	US-09-119-900-8	Sequence 8, Appli
3	1418	100.0	271	30	US-10-453-032-8	Sequence 8, Appli
4	1411	99.5	271	5	US-08-188-721A-2	Sequence 2, Appli
5	1397.5	98.6	272	8	US-08-482-785-15	Sequence 15, Appl
6	1397.5	98.6	272	15	US-09-119-900-15	Sequence 15, Appl
7	1397.5	98.6	272	30	US-10-453-032-15	Sequence 15, Appl
8	1220	86.0	229	8	US-08-482-785-9	Sequence 9, Appli
9	1220	86.0	229	15	US-09-119-900-9	Sequence 9, Appli
10	1220	86.0	229	30	US-10-453-032-9	Sequence 9, Appli
11	371	26.2	303	29	US-10-360-101-263	Sequence 263, App
12	203	14.3	43	8	US-08-482-785-1	Sequence 1, Appli
13	203	14.3	43	15	US-09-119-900-1	Sequence 1, Appli
14	203	14.3	43	30	US-10-453-032-1	Sequence 1, Appli
15	202	14.2	38	8	US-08-482-785-6	Sequence 6, Appli
16	202	14.2	38	15	US-09-119-900-6	Sequence 6, Appli
17	202	14.2	38	30	US-10-453-032-6	Sequence 6, Appli
18	190.5	13.4	252	30	US-09-689-278-2	Sequence 2, Appli
19	182.5	12.9	252	20	US-09-689-278-8	Sequence 8, Appli
20	170	12.0	32	8	US-08-482-785-16	Sequence 16, Appl
21	170	12.0	32	15	US-09-119-900-16	Sequence 16, Appl
22	170	12.0	32	30	US-10-453-032-16	Sequence 16, Appl
23	133	9.4	247	1	PCT-US98-27612-34	Sequence 34, Appl
24	133	9.4	247	16	US-09-221-014-34	Sequence 34, Appl
25	133	9.4	247	19	US-09-561-077C-34	Sequence 34, Appl
26	131	9.2	242	1	PCT-US98-27612-22	Sequence 22, Appl
27	131	9.2	242	16	US-09-221-014-22	Sequence 22, Appl
28	131	9.2	242	19	US-09-561-077C-22	Sequence 22, Appl
29	131	9.2	261	22	US-09-769-736-129	Sequence 129, App
30	131	9.2	261	22	US-09-769-736-129	Sequence 129, App
31	117.5	8.3	274	19	US-09-583-110-3631	Sequence 3631, Ap
32	117.5	8.3	274	31	US-10-640-833-3631	Sequence 3631, Ap
33	113	8.0	97	30	US-10-417-884-5038	Sequence 5038, Ap
34	109.5	7.7	2386	27	US-10-156-761-7751	Sequence 7751, Ap
35	107	7.5	285	1	PCT-US02-36123-3634	Sequence 3634, Ap
36	107	7.5	325	1	PCT-US02-36123-3636	Sequence 3636, Ap
37	106	7.5	23	8	US-08-482-785-4	Sequence 4, Appli
38	106	7.5	23	15	US-09-119-900-4	Sequence 4, Appli
39	106	7.5	23	30	US-10-453-032-4	Sequence 4, Appli
40	106	7.5	1233	20	US-09-602-874C-288	Sequence 288, App
41	106	7.5	1233	20	US-09-605-703B-2698	Sequence 2698, Ap
42	106	7.5	1233	21	US-09-738-626-4312	Sequence 4312, Ap
43	104.5	7.4	455	1	PCT-US97-19575-80	Sequence 80, Appl
44	104.5	7.4	455	1	PCT-US98-22883-82	Sequence 82, Appl
45	104.5	7.4	455	11	US-08-759-739-457	Sequence 457, App

#### ALIGNMENTS

RESULT 1  
US-08-482-785-8  
; Sequence 8, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,785

```
;
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-785-8

Query Match 100.0%; Score 1418; DB 8; Length 271;
Best Local Similarity 100.0%; Pred. No. 5.8e-144;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCRLLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60
Db 1 MNLGSRVFSKCRLLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60
Qy 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
Db 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
Qy 121 SYGVRSQFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180
Db 121 SYGVRSQFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180
Qy 181 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRQGYLYEVAPIYNADELIPRAVVVSMQ 240
Db 181 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRQGYLYEVAPIYNADELIPRAVVVSMQ 240
Qy 241 SSDNTINEKLVYNTANGTYINHGTPQK 271
Db 241 SSDNTINEKLVYNTANGTYINHGTPQK 271

RESULT 2
US-09-119-900-8
; Sequence 8, Application US/09119900
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:

; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-119-900-8

Query Match 100.0%; Score 1418; DB 15; Length 271;
Best Local Similarity 100.0%; Pred. No. 5.8e-144;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MNLGSRVFSKCRLLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60
Db 1 MNLGSRVFSKCRLLVKFSMVALVSATMAVTTVLTALARTQVSNVDVLDGASKY 60
Qy 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
Db 61 LNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
Qy 121 SYGVRSQFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180
Db 121 SYGVRSQFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNA 180
Qy 181 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRQGYLYEVAPIYNADELIPRAVVVSMQ 240
Db 181 VTGTRTONVGRDQKGMRYTEQRAQEWLEARNRQGYLYEVAPIYNADELIPRAVVVSMQ 240
Qy 241 SSDNTINEKLVYNTANGTYINHGTPQK 271
Db 241 SSDNTINEKLVYNTANGTYINHGTPQK 271

RESULT 3
US-10-453-032-8
; Sequence 8, Application US/10453032
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/453,032
; FILING DATE: 03-JUNE-2003
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE: 23-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-453-032-8

Query Match      100.0%; Score 1418; DB 30; Length 271;
Best Local Similarity 100.0%; Pred. No. 5.8e-144;
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVTLENTALARQTQVSNVDVNLNDGASKY 60
DB 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVTLENTALARQTQVSNVDVNLNDGASKY 60
QY 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
DB 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
QY 121 SYGVRSGFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDPMNRSHLIADSLGGDALRVNA 180
DB 121 SYGVRSGFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDPMNRSHLIADSLGGDALRVNA 180
QY 181 VTGTRQNVGGRQKGMRYTEQRAQEWLEANDRGYLYVEVAPIYNADELIPRAVVVSMQ 240
DB 181 VTGTRQNVGGRQKGMRYTEQRAQEWLEANDRGYLYVEVAPIYNADELIPRAVVVSMQ 240
QY 241 SSDNTINEKLVNTANGYTYNHNGTPTQK 271
DB 241 SSDNTINEKLVNTANGYTYNHNGTPTQK 271

RESULT 4
US-08-188-721A-2
; Sequence 2, Application US/08188721A
; GENERAL INFORMATION:
; APPLICANT: Yutsudo, Takashi
; APPLICANT: Okumura, Koichi
; APPLICANT: Iwasaki, Makoto
; APPLICANT: Hara, Ayako
; APPLICANT: Kishishita, Masamichi
; APPLICANT: Takeda, Yoshifumi
; APPLICANT: Igarashi, Hiesanaga
; APPLICANT: Hinuma, Yorio
; TITLE OF INVENTION: Mitogenic Factor, Gene Thereof and
; TITLE OF INVENTION: Method of Microdetection Therefor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,721A
; FILING DATE: 31-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977

; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-785-8.rapm

REFERENCE/DOCKET NUMBER: 1422-178P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-721A-2

Query Match      99.5%; Score 1411; DB 5; Length 271;
Best Local Similarity 99.3%; Pred. No. 3.3e-143;
Matches 269; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVTLENTALARQTQVSNVDVNLNDGASKY 60
DB 1 MNLGSRVFSKCRVLKFSMVALVSATMAVTTVTLENTALARQTQVSNVDVNLNDGASKY 60
QY 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
DB 61 LNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEG 120
QY 121 SYGVRSGFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDPMNRSHLIADSLGGDALRVNA 180
DB 121 SYGVRSGFGKQNPAGWTGNPNHVKYKIEWLNGLSYVGDPMNRSHLIADSLGGDALRVNA 180
QY 181 VTGTRQNVGGRQKGMRYTEQRAQEWLEANDRGYLYVEVAPIYNADELIPRAVVVSMQ 240
DB 181 VTGTRQNVGGRQKGMRYTEQRAQEWLEANDRGYLYVEVAPIYNADELIPRAVVVSMQ 240
QY 241 SSDNTINEKLVNTANGYTYNHNGTPTQK 271
DB 241 SSDNTINEKLVNTANGYTYNHNGTPTQK 271

RESULT 5
US-08-482-785-15
; Sequence 15, Application US/08482785
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,785
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
```

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-785-15

Query Match 98.6%; Score 1397.5; DB 8; Length 272;

Best Local Similarity 99.6%; Pred. No. 9.7e-142;

Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	2	NLGSRVFSKCKRLVKFSMVALVSATMAVTTVLENTALARTQVSNVDVNDGASKYL	61
DB	4	NLGSRVFSKCKRLVKFSMVALVSATMAVTTVLENTALA-QTVSNVDVNDGASKYL	62
QY	62	NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS	121
DB	63	NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS	122
QY	122	YGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSVVGDFWNRSHLIADSLGGDALRVNAV	181
DB	123	YGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSVVGDFWNRSHLIADSLGGDALRVNAV	182
QY	182	TGTRTQNVGGDRQGGMRYTEQRAQEWLEARNRGYLYEVAPIYNADELIPRAVVVSMQS	241
DB	183	TGTRTQNVGGDRQGGMRYTEQRAQEWLEARNRGYLYEVAPIYNADELIPRAVVVSMQS	242
QY	242	SDNTINEKVLVNTANGYTYNHNGTPTQK	271
DB	243	SDNTINEKVLVNTANGYTYNHNGTPTQK	272

## RESULT 6

US-09-119-900-15

Sequence 15, Application US/09119900

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/119,900

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/082,845

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.

REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:  
LENGTH: 272 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-119-900-15

Query Match 98.6%; Score 1397.5; DB 15; Length 272;

Best Local Similarity 99.6%; Pred. No. 9.7e-142;

Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	2	NLGSRVFSKCKRLVKFSMVALVSATMAVTTVLENTALARTQVSNVDVNDGASKYL	61
DB	4	NLGSRVFSKCKRLVKFSMVALVSATMAVTTVLENTALA-QTVSNVDVNDGASKYL	62
QY	62	NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS	121
DB	63	NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSKLDLGRTRTARGTLTYANVEGS	122
QY	122	YGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSVVGDFWNRSHLIADSLGGDALRVNAV	181
DB	123	YGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSVVGDFWNRSHLIADSLGGDALRVNAV	182
QY	182	TGTRTQNVGGDRQGGMRYTEQRAQEWLEARNRGYLYEVAPIYNADELIPRAVVVSMQS	241
DB	183	TGTRTQNVGGDRQGGMRYTEQRAQEWLEARNRGYLYEVAPIYNADELIPRAVVVSMQS	242
QY	242	SDNTINEKVLVNTANGYTYNHNGTPTQK	271
DB	243	SDNTINEKVLVNTANGYTYNHNGTPTQK	272

## RESULT 7

US-10-453-032-15

Sequence 15, Application US/10453032

GENERAL INFORMATION:

APPLICANT: Adams, Craig W.

APPLICANT: Pang, Patty P.-Y.

APPLICANT: Belei, Marina

TITLE OF INVENTION: Recombinant DNase B Derived from

TITLE OF INVENTION: Streptococcus pyogenes

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak

STREET: 225 South Lake Avenue, Ninth Floor

CITY: Pasadena

STATE: California

COUNTRY: USA

ZIP: 91001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/453,032

FILING DATE: 03-JUNE-2003

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/082,845

FILING DATE: 23-JUNE-1993

ATTORNEY/AGENT INFORMATION:

NAME: Farber, Michael B.

REGISTRATION NUMBER: 32,612

REFERENCE/DOCKET NUMBER: 9521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 272 amino acids

TYPE: amino acid

TOPOLOGY: linear



```
; MOLECULE TYPE: protein
US-10-453-032-15
Query Match      98.6%; Score 1397.5; DB 30; Length 272;
Best Local Similarity 99.6%; Pred. No. 9.7e-142;
Matches 269; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2 NLLGSRVFSKCLVFKFMSVALVSATMAVTTVLTALARTQVSDVNDGASKYL 61
DB 4 NLLGSRVFSKCLVFKFMSVALVSATMAVTTVLTALARTQVSDVNDGASKYL 62
QY 62 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDELGRTRTARGTLTVANVEGS 121
DB 63 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDELGRTRTARGTLTVANVEGS 122
QY 122 YGVRSQSGKQNPAGWTGNPNHVKYKIEWNLGLSYVGD FWNRSHLIADSLGGDALRVNAV 181
DB 123 YGVRSQSGKQNPAGWTGNPNHVKYKIEWNLGLSYVGD FWNRSHLIADSLGGDALRVNAV 182
QY 182 TGTTRQNVGGDRQGGHRYTEQRAQEWLEARNRDGILYEVAPIYNADDELIPRAVVVSMQS 241
DB 183 TGTTRQNVGGDRQGGHRYTEQRAQEWLEARNRDGILYEVAPIYNADDELIPRAVVVSMQS 242
QY 242 SDNTINEKLVYNTANGTYTINHGTPQK 271
DB 243 SDNTINEKLVYNTANGTYTINHGTPQK 272

RESULT 8
US-08-482-785-9
; Sequence 9, Application US/08482785
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,785
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
; Query Match      86.0%; Score 1220; DB 15; Length 229;
; Best Local Similarity 100.0%; Pred. No. 1.2e-122;
; Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ROTQVSNVVDNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDE 102
DB 1 ROTQVSNVVDNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDE 60
QY 103 LGRTRTARGTLTVANVEGSYGVRSQSGKQNPAGWTGNPNHVKYKIEWNLGLSYVGD FWN 162
DB 61 LGRTRTARGTLTVANVEGSYGVRSQSGKQNPAGWTGNPNHVKYKIEWNLGLSYVGD FWN 120
QY 163 RSHLIADSLGGDALRVNAVTVGTTRQNVGGDRQGGHRYTEQRAQEWLEARNRDGILYEV 222
DB 121 RSHLIADSLGGDALRVNAVTVGTTRQNVGGDRQGGHRYTEQRAQEWLEARNRDGILYEV 180
QY 223 PIYNADDELIPRAVVVSMQSSDNTINEKLVYNTANGTYTINHGTPQK 271
DB 181 PIYNADDELIPRAVVVSMQSSDNTINEKLVYNTANGTYTINHGTPQK 229

RESULT 9
US-09-119-900-9
; Sequence 9, Application US/09119900
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,900
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 229 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pyogenes
;
; US-09-119-900-9
;
; Query Match      86.0%; Score 1220; DB 15; Length 229;
```

Best Local Similarity 100.0%; Pred. No. 1.2e-122; Indels 0; Gaps 0;  
Matches 229; Conservative 0; Mismatches 0;

Qy 43 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALFPKAGDILYSKLE 102  
|||||  
Db 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALFPKAGDILYSKLE 60  
|||||

Qy 103 LGRTRTARGTLTYANVEGSGVROSGKQNPAGWTGNPNHVYKIEWLNGLSVGDFWN 162  
|||||  
Db 61 LGRTRTARGTLTYANVEGSGVROSGKQNPAGWTGNPNHVYKIEWLNGLSVGDFWN 120  
|||||

Qy 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 222  
|||||  
Db 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 180  
|||||

Qy 223 PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYINYHNGTPTQK 271  
|||||  
Db 181 PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYINYHNGTPTQK 229  
|||||

## RESULT 10

US-10-453-032-9  
; Sequence 9, Application US/10453032  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10453,032  
; FILING DATE: 03-JUNE-2003  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: 23-JUNE-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
US-10-453-032-9

Query Match 86.0%; Score 1220; DB 30; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.2e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALFPKAGDILYSKLE 102  
|||||

Db 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALFPKAGDILYSKLE 60  
|||||

Qy 103 LGRTRTARGTLTYANVEGSGVROSGKQNPAGWTGNPNHVYKIEWLNGLSVGDFWN 162  
|||||

Db 61 LGRTRTARGTLTYANVEGSGVROSGKQNPAGWTGNPNHVYKIEWLNGLSVGDFWN 120  
|||||

Qy 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 222  
|||||

Db 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVA 180  
|||||

Qy 223 PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYINYHNGTPTQK 271  
|||||

Db 181 PIYNADELIPRAVVVSMOSSDNTINEKVLVNTANGTYINYHNGTPTQK 229  
|||||

## RESULT 11

US-10-360-101-263  
; Sequence 263, Application US/10360101  
; GENERAL INFORMATION:  
; APPLICANT: Moll, Gert N.  
; APPLICANT: Leenhouts, Cornelis J.  
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way  
; FILE REFERENCE: 2183-5673  
; CURRENT APPLICATION NUMBER: US/10/360,101  
; CURRENT FILING DATE: 2003-02-07  
; PRIOR APPLICATION NUMBER: EP 02077060.8  
; PRIOR FILING DATE: 2002-05-24  
; NUMBER OF SEQ ID NOS: 309  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 263  
; TYPE: PRT  
; LENGTH: 303  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: sequence of streptodornase  
US-10-360-101-263

Query Match 26.2%; Score 371; DB 29; Length 303;  
Best Local Similarity 32.6%; Pred. No. 1.3e-30;  
Matches 89; Conservative 36; Mismatches 84; Indels 64; Gaps 5;

Qy 38 NTALARQTVSNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTQITPALFPK----- 91  
|||||

Db 7 NTVLAKTVSVNQ-----TYGEYKDYTYVIGESNIDQSAFPKYYKTE 48  
|||||

Qy 92 -----AGDILYSKLELGRTRTARGTLTYANVEGSGVROSGKQNPAGW 137  
|||||

Db 49 RVYKQGTSEKRVTVSDVYNPLDGYKRTGAYGVVTKMDMSKGYREKWEWNPESGW 108  
|||||

Qy 138 -----TGN-----PNHVYKIEWLNGLSVGDFWNRSHLTDLSGG 173  
|||||

Db 109 PRFYNRADNEEISEKEYDSRRTKSYKVTNNVPVVLTLTKGKYNHSLFVASHLPADSLGG 168  
|||||

Qy 174 DALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYEVAIYNADELIPR 233  
:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 169 KSIKNAITGTMQNVGTR--KGMQYIEKKVLSHITKNPDVYVYSAIPEYQGAELLAR 226  
:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Qy 234 AVVVSQSSDNTINEKVLVNTANGTYINYHNG 266  
:|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

Db 227 SVLVSSDSDGVNETVRVNTADGFNINTEKG 259  
:|||:|||||:|||||:|||||:|||||:|||||:|||||

## RESULT 12

US-08-482-785-1  
; Sequence 1, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,785  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE: 24-FEB-1995  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-08-482-785-1  
Query Match 14.3%; Score 203; DB 8; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43  
Db 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43  
RESULT 13  
US-09-119-900-1  
Sequence 1, Application US/09119900  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-1  
Query Match 14.3%; Score 203; DB 15; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43  
Db 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43  
RESULT 14  
US-10-453-032-1  
Sequence 1, Application US/10453032  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,032  
FILING DATE: 03-JUNE-2003  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: 23-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 43 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

Search completed: January 5, 2004, 18:55:06  
Job time : 325.403 secs

MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-10-453-032-1

Query Match 14.3%; Score 203; DB 30; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNLGSRVFSKCKELVKFMSVALVSATMAVTVTTLTALAR 43  
|||||  
DB 1 MNLGSRVFSKCKELVKFMSVALVSATMAVTVTTLTALAR 43

RESULT 15  
US-08-482-785-6  
; Sequence 6, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,785  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,889  
; FILING DATE: 24-FEB-1995  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
US-08-482-785-6

Query Match 14.2%; Score 202; DB 8; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ROTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
|||||  
DB 1 ROTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 38

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 5, 2004, 18:39:19 ; Search time 42.9208 Seconds  
(without alignments)  
408.499 Million cell updates/sec

Title: US-08-482-785-8  
Perfect score: 1418  
Sequence: 1 MNLLGRRVSKCRLVKFS.....VYNTANGTYINHGTFQK 271

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 320769 seqs, 64697744 residues

Total number of hits satisfying chosen parameters: 320769

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents, AA New:  
1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pap.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pap.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pap.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pap.\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pap.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pap.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1418	100.0	271	4	US-08-472-630-8
2	1397.5	98.6	272	4	US-08-472-630-15
3	1220	86.0	229	4	US-08-472-630-9
4	203	14.3	43	4	US-08-472-630-1
5	202	14.2	38	4	US-08-472-630-6
6	170	12.0	32	4	US-08-472-630-16
7	126.5	8.9	274	1	PCT-US03-27401-436
8	126.5	8.9	274	6	US-10-472-928-4092
9	106	7.5	23	4	US-08-472-630-4
10	97	6.8	281	6	US-10-425-114A-51758
11	92	6.5	446	6	US-10-425-114A-57900
12	92	6.5	450	6	US-10-425-114A-56379
13	90	6.3	398	5	US-09-897-516A-6933
14	89	6.3	396	6	US-10-679-063-26323
15	88	6.2	646	1	PCT-US03-27401-377
16	88	6.2	646	6	US-10-472-928-2716
17	88	6.2	646	6	US-10-474-776-328
18	87.5	6.2	608	6	US-10-679-063-21735
19	87.5	6.2	1738	1	PCT-US03-28227-4134
20	87	6.1	362	6	US-10-250-682-2
21	87	6.1	428	1	PCT-US02-34769-8
22	87	6.1	429	1	PCT-US02-34769-16
23	87	6.1	430	5	US-09-614-150A-25899
24	86.5	6.1	280	6	US-10-425-114A-52835
25	86.5	6.1	329	6	US-10-425-114A-69718
26	86.5	6.1	331	6	US-10-425-114A-53765

27	86.5	6.1	566	6	US-10-425-114A-47841	Sequence 47841, A
28	86.5	6.1	1709	7	US-60-490-890-1987	Sequence 1987, Ap
29	86	6.1	1391	6	US-10-687-046-11	Sequence 11, Appl
30	86	6.1	1391	6	US-10-687-046-15	Sequence 15, Appl
31	85.5	6.0	482	5	US-09-897-516A-6199	Sequence 6199, Ap
32	85.5	6.0	1475	1	PCT-US03-06962-34	Sequence 34, Appl
33	84.5	6.0	818	5	US-09-897-516A-5123	Sequence 5123, Ap
34	84.5	6.0	1436	6	US-10-687-046-13	Sequence 13, Appl
35	84.5	6.0	5635	6	US-10-451-168-78	Sequence 78, Appl
36	83.5	5.9	1336	6	US-10-019-065A-33	Sequence 33, Appl
37	83	5.9	355	1	PCT-US03-19153-164	Sequence 164, App
38	83	5.9	355	6	US-10-463-720-164	Sequence 164, App
39	83	5.9	705	1	PCT-US03-35733-25	Sequence 25, Appl
40	83	5.9	876	1	PCT-US03-35733-31	Sequence 31, Appl
41	83	5.9	1203	6	US-10-472-928-4086	Sequence 4086, Ap
42	82.5	5.8	1795	4	US-08-973-363A-17	Sequence 17, Appl
43	82.5	5.8	2537	1	PCT-US03-30720-720	Sequence 720, App
44	82	5.8	285	6	US-10-009-384-30	Sequence 30, Appl
45	82	5.8	540	5	US-09-674-546A-2941	Sequence 2941, Ap

## ALIGNMENTS

### RESULT 1

US-08-472-630-8  
; Sequence 8, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patsy P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-472-630-8

Query Match 100.0%; Score 1418; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 66-106;  
Matches 271; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNLLGSRRVFSKCRLLVFSFMSVALVSATMATVTTVLTENTALARQTQVSNDVNLDCASKY	60
Db	1	MNLLGSRRVFSKCRLLVFSFMSVALVSATMATVTTVLTENTALARQTQVSNDVNLDCASKY	60
Qy	61	LNELALAWTENDSPNYKYKLTGTSQITPALFPKAGDILYSKLDELGRTRTARGTLTYANVEG	120
Db	61	LNELALAWTENDSPNYKYKLTGTSQITPALFPKAGDILYSKLDELGRTRTARGTLTYANVEG	120
Qy	121	SYGVROSFGKNQPAWGTPGNPHNVKYLEWINGLSYVGDFWNRSHLIADSLGGDALRVNA	180
Db	121	SYGVROSFGKNQPAWGTPGNPHNVKYLEWINGLSYVGDFWNRSHLIADSLGGDALRVNA	180
Qy	181	VTGTRTQNVGGRDQGGWMRYTEORAQEWLEANRQGYLYYEVAPIYNADELI	240
Db	181	VTGTRTQNVGGRDQGGWMRYTEORAQEWLEANRQGYLYYEVAPIYNADELI	240
Qy	241	SSDNTINEKVLVYNTANGYTYINHGNTPTQK	271
Db	241	SSDNTINEKVLVYNTANGYTYINHGNTPTQK	271

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

62 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDELGRTRTARGTLTYANVEGS 121  
63 NEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYKSLDELGRTRTARGTLTYANVEGS 122  
122 YGVRQSEFGKNONPAGWTCGNPHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNAV 181  
123 YGVRQSEFGKNONPAGWTCGNPHVKYKIEWLNGLSYVGDFWNRSHLIADSLGGDALRVNAV 182  
182 TGTRTQNVGGRDQKGGWRYTEQRAQEWLEARNRDCGLYVEVAPIYNADELIPRAVVVSMQS 241  
183 TGTRTQNVGGRDQKGGWRYTEQRAQEWLEARNRDCGLYVEVAPIYNADELIPRAVVVSMQS 242  
242 SDNTINEKVLVYNTANGYTYINHNGTPTQK 271  
243 SDNTINEKVLVYNTANGYTYINHNGTPTQK 272

RESULT 3  
US-08-472-630-9  
; Sequence 9, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belel, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 229 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:  
YS-08-472-630-9

Qy 103 LGRTTRTARGTLTYANVGGSYGVROSPGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 162  
Db 61 LGRTTRTARGTLTYANVGGSYGVROSPGKQNPAGTGNPNHVKYKIEWLNGLSVVGDFWN 120  
Qy 163 RSHLIADSLGDLARLVNAVVTGTRTQNVGGDKGGRYTBQRAQEWLEARNRDLGYLYVEVA 222  
Db 121 RSHLIADSLGDLARLVNAVVTGTRTQNVGGDKGGRYTBQRAQEWLEARNRDLGYLYVEVA 180  
Qy 223 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTTANGTYTINYHNGTPTQK 271  
Db 181 PIYNADELIPRAVVVSMQSSDNTINEKVLVYNTTANGTYTINYHNGTPTQK 229

## RESULT 4

US-08-472-630-1  
; Sequence 1, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 43 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-08-472-630-1

Query Match 14.3%; Score 203; DB 4; Length 43;  
Best Local Similarity 100.0%; Pred. No. 7.4e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43  
Db 1 MNLGSRVFSKKRLVKFSWALVSATMAVTTVLTALAR 43

## RESULT 5

US-08-472-630-6  
; Sequence 6, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-08-472-630-6  
Query Match 14.2%; Score 202; DB 4; Length 38;  
Best Local Similarity 100.0%; Pred. No. 7.8e-10;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 43 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 80  
Db 1 RQTQVSNVDVLDGASKYLNEALAWTFNDSPNYKTLG 38  
RESULT 6  
US-08-472-630-16  
; Sequence 16, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA

ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,630  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/092,845  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-08-472-630-16

```

RESULT 7
PCT-US03-27401-436
; Sequence 436, Application PC/TUS0327401
; GENERAL INFORMATION:
; APPLICANT: TUFTS UNIVERSITY
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF ACTIVE INFECTION
; FILE REFERENCE: 700355-52941-PCT
; CURRENT APPLICATION NUMBER: PCT/US03/27401
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/407,082
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 436
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
PCT-US03-27401-436

```

Query Match	8.9%;	Score 126.5;	DB 1;	Length 274;
Best Local Similarity	21.9%;	Pred. No. 0.0061;		
Matches	59;	Conservative 37;	Mismatches 108;	Indels 65; Gaps 11
Qy	32	TTVLTLENTARQTVSDNVLDNGASKYLNEALAWTFND-----	71	
Db	37	TNLSQKQKASEAPDSQALAESVTTDAVKSQIKGSLEWNGCAFTVNGKTNLDAKVSSKEPY	96	
Qy	72	SPNYKTKTQSTQITPALPKAGDILYKSLDE--LGRTRTARGTLTVANVEGSGVGRQSEF	129	
Db	97	ADNKTITVQKEIV-----PTVANALLSKATPQYKNRKETGNGSTSW-----	137	

```

130 QY KNQNPAGWTGPNHVKYKTEWLNGLSYVCGDFWNRGHLIADSL-----VNAV 181
138 Db ---TPPGW-----HQVKNLKG-SYTHAV-DRGHLLGYALIGLGGFDASTSNPKNIA 184
182 QY TGTRQTVNGVRDQKGMRYTEQRAQEWLEFANRDGVLYYEVAFYINADE-LIPRAVVVSMQ 240
185 Db VTAWANQAQAEVSTGQNYYESKVRKALDQNK-RVRYRVTLYYASNEDLVPASQIEAK 242
241 QY SSDNTINEKVLVYNTANGTYTINYHNGTPT 269
243 Db SSDGELEFNVLVPNVQKGLQLDYRTGEVT 271

RESULT 8
US-10-472-928-4092
; Sequence 4092, Application US/10472928
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4092
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: DNA-entry nuclease (enda)
; OTHER INFORMATION: Cellular location: outside
; OTHER INFORMATION: Similar to strain R6 sequence 15903821 (e-156)
US-10-472-928-4092

```

```

RESULT 9
US-08-472-630-4
; Sequence 4, Application US/08472630
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; Pang, Patty P.-Y.
; Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
;
;
;

```



ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
City: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,630  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-472-630-4

Query Match 7.5%; Score 106; DB 4; Length 23;  
Best Local Similarity 95.7%; Pred. No. 0.021;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 44 QTQVSNDDVNLNDGASKYLNEALA 66  
Db 1 QTQVSNDDVNLNDGASKYLNEALA 23

RESULT 10  
US-10-425-114A-51758  
; Sequence 51758, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 51758  
; LENGTH: 281  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700869544\_FLI.pep  
US-10-425-114A-51758

Query Match 6.8%; Score 97; DB 6; Length 281;  
Best Local Similarity 23.2%; Pred. No. 1.4;

Matches 51; Conservative 39; Mismatches 74; Indels 56; Gaps 12;  
Qy 56 GASKYLNEALAWTFNDSPNYKTL-GTSQITPALFPKAGDILYKLDLDELGRTRTARGTLT 114  
Db 6 GAPTVLNMII-----NSSPKVQKPLPGKVQVMTGAPPPDVIF-RMBELG----- 50  
Qy 115 YANVEGSGYVQSGKQNPAG-----WTGNPNHVYKIEWLNGLSYVG-----DF 160  
Db 51 -FNVTHSYGLTETEG-----PASICTKPEWDNLPQDAQAKLKARQGVAVHVGMEGLDVKDP 105  
Qy 161 WNRSHLIAD--SLGGDALRVNAVGTGRTONTVGRDQ--KGGMRYTEQRAQEWLEANRDGY 216  
Db 106 HTMKSVFADAKTGMVFRGTVWNGYLKDLKATQEAFFKGGFWTGDLCVK-----HPDCY 161  
Qy 217 LYEVAVPIYNADELIPRA--VWVSMQSSDNTINEKVLVYN 254  
Db 162 I-----ELKDRSKDIIISGGENISTIELEGVIFS 190

## RESULT 11

US-10-425-114A-57900  
; Sequence 57900, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 57900  
; LENGTH: 446  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: UC-ZMFLM017008B12\_FLI.pep  
US-10-425-114A-57900

Query Match 6.5%; Score 92; DB 6; Length 446;  
Best Local Similarity 20.6%; Pred. No. 5.7;  
Matches 62; Conservative 37; Mismatches 92; Indels 110; Gaps 14;

Qy 38 NTALARQTQVSNDDVNLNDGASK-----YLNALAWTFNDSPNYKTLGTSQITPALFPKAG 93  
Db 108 NENLIKQTA--DALVNTGLAKLGYEYVNIIDCWAESDRDYQGSFVANRQTFFSGIKALA 164  
Qy 94 DILYSKLDELG-----RRTAR-----GTLTYANVE-----CSYGV----- 124  
Db 165 DYVHAKGLKGIYSDAGTRTCQKMPGSLDHEEQDKVTFFSSWGIDYLYKYNCDNAGRSVM 224  
Qy 125 -----RQSGK-----NONPAGTGNPNHVYKIEWLNGLSVVGDFWNRSH 165  
Db 225 ERTKSNAMKTYGKGIFFSLCEWGRQNPATWAG-----MGNSWRTTD 268  
Qy 166 LIADSLGGDALRVNAVGTGRTON-----VGGRDQ-----KGGMRYTEQRAQEWLE 210  
Db 269 DIADNWG-----SMTSRADQNRWASVAGPGWNDPDMLEVNGCGMSEAYRS----- 316  
Qy 211 ANRDGYLYEVAVPIYNADELIPRAV-VWVSMQSSDNTINEKVLVYNT-----ANGTYINHN 265  
Db 317 -----HFSIWALAKAPLIIGCDVRAMSQQTGILSNSEVIAVNDQSQAGKKVQSSN 369  
Qy 266 G 266  
Db 370 G 370

## RESULT 12

```

US-10-425-114A-56379
; Sequence 56379, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56379
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73183G03_FLI.pep
US-10-425-114A-56379

Query Match          6.5%; Score 92; DB 6; Length 450;
Best Local Similarity 20.6%; Pred. No. 5.7;
Matches 62; Conservative 37; Mismatches 92; Indels 110; Gaps 14;

Qy 38 NTALARQTVSNDDVVLNDGASK-----YLNEALAWTFNDSPNYKTLGTQITPALFPKAG 93
Db 112 NENLIKOTA---DALVNTGLAKLGYEVYNIIDCWAEISDRDYQGSFVANRQTFFPSGKALA 168
Qy 94 DILYSKLDLGL-----RRTAR-----GTLTYANVE-----GSYGV----- 124
Db 169 DYVHAKGLKGIYSDAGTRTCSQKMPGSLDHEEQDVKTFFSWGIDLYKYDNCNDAGRSVM 228
Qy 125 -----RSGFK-----NONPAGWTGNPNHVKYKIEWLNGLSYVGDFWNRSH 165
Db 229 ERYTKMSNMATYKGIFFSLCEWGRQNPATWAG-----MGNSWRTTD 272
Qy 166 LIADSLGDLRVNAVTCRTQN-----VGRDQ-----KGGMRYTEQRAQEWLE 210
Db 273 DIADNWG-----SMTSRAQNDRWASVAGPGGWNDPDMLEVGNGMSEAEYRS----- 320
Qy 211 ANRDGILYVEVAPTYNADLIPRAV-VVSMOSSDNTINEKVLVYNT---ANGYTYNYHN 265
Db 321 -----HFSWTALAKAPLLIGCDVRAMSQQTWGLSLNSEVIAVNVQDSQAGGKKVQSSN 373
Qy 266 G 266
Db 374 G 374

RESULT 13
US-09-897-516A-6933
; Sequence 6933, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 6933
; LENGTH: 398

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 5, 2004, 18:38:21 ; Search time 273.282 Seconds  
(without alignments)  
762.478 Million cell updates/sec

Title: US-08-482-785-9  
Perfect score: 1220  
Sequence: 1 RQTVQNDVVLDGASKYLN.....VYNTANGYTYNHNGTPTQK 229

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.\*  
1: /cgn2\_6/ptodata/2/paa/PCTUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US096\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US097A\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US097B\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US099A\_COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US099B\_COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*  
27: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*  
28: /cgn2\_6/ptodata/2/paa/US102\_COMB.pep.\*  
29: /cgn2\_6/ptodata/2/paa/US103\_COMB.pep.\*  
30: /cgn2\_6/ptodata/2/paa/US104\_COMB.pep.\*  
31: /cgn2\_6/ptodata/2/paa/US106\_COMB.pep.\*  
32: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1220	100.0	229	8	US-08-482-785-9 Sequence 9, Appli

ALIGNMENTS

RESULT 1  
US-08-482-785-9  
; Sequence 9, Application US/08482785  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; TITLE OF INVENTION: Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,785

2	1220	100.0	229	15	US-09-119-900-9	Sequence 9, Appli
3	1220	100.0	229	30	US-10-453-032-9	Sequence 9, Appli
4	1220	100.0	271	8	US-08-482-785-8	Sequence 8, Appli
5	1220	100.0	271	15	US-09-119-900-8	Sequence 8, Appli
6	1220	100.0	271	30	US-10-453-032-8	Sequence 8, Appli
7	1215	99.6	272	8	US-08-482-785-15	Sequence 15, Appl
8	1215	99.6	272	15	US-09-119-900-15	Sequence 15, Appl
9	1215	99.6	272	30	US-10-453-032-15	Sequence 15, Appl
10	1213	99.4	271	5	US-08-188-721A-2	Sequence 2, Appli
11	365	29.9	303	29	US-10-360-101-263	Sequence 263, App
12	202	16.6	38	8	US-08-482-785-6	Sequence 6, Appli
13	202	16.6	38	15	US-09-119-900-6	Sequence 6, Appli
14	202	16.6	38	30	US-10-453-032-6	Sequence 6, Appli
15	179.5	14.7	252	20	US-09-689-278-2	Sequence 2, Appli
16	171.5	14.1	252	20	US-09-689-278-8	Sequence 8, Appli
17	170	13.9	32	8	US-08-482-785-16	Sequence 16, Appl
18	170	13.9	32	15	US-09-119-900-16	Sequence 16, Appl
19	170	13.9	32	30	US-10-453-032-16	Sequence 16, Appl
20	131	10.7	261	22	US-09-769-736-129	Sequence 129, App
21	126.5	10.4	247	1	PCT-US98-27612-34	Sequence 34, Appl
22	126.5	10.4	247	16	US-09-221-014-34	Sequence 34, Appl
23	126.5	10.4	247	19	US-09-561-077C-34	Sequence 34, Appl
24	124.5	10.2	242	1	PCT-US98-27612-22	Sequence 22, Appl
25	124.5	10.2	242	16	US-09-221-014-22	Sequence 22, Appl
26	124.5	10.2	242	19	US-09-561-077C-22	Sequence 22, Appl
27	122.5	10.0	274	22	US-09-769-744A-168	Sequence 168, App
28	113.5	9.3	274	19	US-09-583-110-3631	Sequence 3631, Ap
29	113.5	9.3	274	31	US-10-640-833-3631	Sequence 3631, Ap
30	113	9.3	97	30	US-10-417-884-5038	Sequence 5038, Ap
31	107	8.8	285	1	PCT-US02-36123-3634	Sequence 3634, Ap
32	107	8.8	325	1	PCT-US02-36123-3636	Sequence 3636, Ap
33	106	8.7	23	8	US-08-482-785-4	Sequence 4, Appli
34	106	8.7	23	15	US-09-119-900-4	Sequence 4, Appli
35	106	8.7	23	30	US-10-453-032-4	Sequence 4, Appli
36	106	8.7	1233	20	US-09-602-874C-288	Sequence 288, App
37	106	8.7	1233	20	US-09-605-703B-2698	Sequence 2698, App
38	106	8.7	1233	21	US-09-738-626-4312	Sequence 4312, Ap
39	104.5	8.6	455	1	PCT-US97-19575-80	Sequence 80, Appl
40	104.5	8.6	455	1	PCT-US98-22883-82	Sequence 82, Appl
41	104.5	8.6	455	11	US-08-759-739-457	Sequence 457, App
42	104.5	8.6	455	13	US-08-993-001-82	Sequence 82, Appl
43	104.5	8.6	455	13	US-08-993-002A-4969	Sequence 4969, Ap
44	104.5	8.6	455	29	US-10-335-977-4969	Sequence 4969, Ap
45	104.5	8.6	486	11	US-08-759-739-359	Sequence 359, App



REFERENCE/DOCKET NUMBER: 9521  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 229 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-10-453-032-9

Query Match 100.0%; Score 1220; DB 30; Length 229;  
Best Local Similarity 100.0%; Pred. No. 1.9e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSGITPALPKAGDILYSLKDE 60  
DB 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSGITPALPKAGDILYSLKDE 60  
QY 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 120  
DB 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 120  
QY 121 RSHLIADSLGDDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
DB 121 RSHLIADSLGDDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
QY 181 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGTYINHNGTPTQK 229  
DB 181 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGTYINHNGTPTQK 229

## RESULT 4

US-08-482-785-8  
Sequence 8, Application US/08482785  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,785  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,889  
FILING DATE: 24-FEB-1995  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELEPHONE: (818) 796-4000

TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-482-785-8

Query Match 100.0%; Score 1220; DB 8; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.4e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSGITPALPKAGDILYSLKDE 60  
DB 43 ROTQVNDVVLNDGASKYLNEALAWTFNDSPNYKTLGTSGITPALPKAGDILYSLKDE 102  
QY 61 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 120  
DB 103 LGRTTRTARGTLTYANVEGSGVRSFGKQNPAGTGNPNHVKYKIEWLNGLSYVGDFWN 162  
QY 121 RSHLIADSLGDDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 180  
DB 163 RSHLIADSLGDDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGYLYYEVA 222  
QY 181 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGTYINHNGTPTQK 229  
DB 223 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGTYINHNGTPTQK 271

## RESULT 5

US-09-119-900-8  
Sequence 8, Application US/09119900  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 271 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-119-900-8

Query Match 100.0%; Score 1220; DB 15; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.4e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQVNDVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDE 60  
Db 43 RQTQVNDVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDE 102

Qy 61 LGRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 120  
Db 103 LGRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 162

Qy 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNR DGYLYYEVA 180  
Db 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNR DGYLYYEVA 222

Qy 181 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGYTYNHNGTPTQK 229  
Db 223 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGYTYNHNGTPTQK 271

## RESULT 6

US-10-453-032-8

; Sequence 8, Application US/10453032

; GENERAL INFORMATION:

; APPLICANT: Adams, Craig W.

; APPLICANT: Pang, Patty P.-Y.

; APPLICANT: Belei, Marina

; TITLE OF INVENTION: Recombinant DNase B Derived from

; TITLE OF INVENTION: Streptococcus pyogenes

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheldon &amp; Mak

; STREET: 225 South Lake Avenue, Ninth Floor

; CITY: Pasadena

; STATE: California

; COUNTRY: USA

; ZIP: 91001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/453,032

; FILING DATE: 03-JUNE-2003

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/082,845

; FILING DATE: 23-JUNE-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Farber, Michael B.

; REGISTRATION NUMBER: 32,612

; REFERENCE/DOCKET NUMBER: 9521

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (818) 796-4000

; TELEFAX: (818) 795-6321

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 271 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-10-453-032-8

Query Match 100.0%; Score 1220; DB 30; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.4e-122;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQVNDVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDE 60  
Db 43 RQTQVNDVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDE 102

Qy 61 LGRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 120  
Db 103 LGRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 162

Qy 121 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNR DGYLYYEVA 180  
Db 163 RSHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNR DGYLYYEVA 222

Qy 181 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGYTYNHNGTPTQK 229  
Db 223 PIYNADLIPRAVVVSMQSSDNTINEKLVYNTANGYTYNHNGTPTQK 271

## RESULT 7

US-08-482-785-15

; Sequence 15, Application US/08482785

; GENERAL INFORMATION:

; APPLICANT: Adams, Craig W.

; APPLICANT: Pang, Patty P.-Y.

; APPLICANT: Belei, Marina

; TITLE OF INVENTION: Recombinant DNase B Derived from

; TITLE OF INVENTION: Streptococcus pyogenes

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheldon &amp; Mak

; STREET: 225 South Lake Avenue, Ninth Floor

; CITY: Pasadena

; STATE: California

; COUNTRY: USA

; ZIP: 91001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/482,785

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/393,889

; FILING DATE: 24-FEB-1995

; APPLICATION NUMBER: US/08/082,845

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Farber, Michael B.

; REGISTRATION NUMBER: 32,612

; REFERENCE/DOCKET NUMBER: 9521

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (818) 796-4000

; TELEFAX: (818) 795-6321

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 272 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-482-785-15

Query Match 99.6%; Score 1215; DB 8; Length 272;  
Best Local Similarity 100.0%; Pred. No. 8.5e-122;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTVQSNVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDEL 61  
Db 45 QTVQSNVLDGASKYLNEALAWTFNDSNNYKTLGTSQITPALPPKAGDILYSKLDEL 104

Qy 62 GRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 121  
Db 105 GRTRTARGTLTYANVEGSGVRSFGKQNPAGWTGNPNHVYKIEWLNGLSVVGD FWN 164

Qy 122 SHLIADSLGGDALRVNAVVTGTRTQNVGGRDQKGMRYTEQRAQEWLEARNR DGYLYYEVA 181



Db 165 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGILYYEVAP 224  
Qy 182 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 229  
Db 225 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 272

RESULT 8  
US-09-119-900-15  
; Sequence 15, Application US/09119900  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,900  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-119-900-15

Query Match 99.6%; Score 1215; DB 15; Length 272;  
Best Local Similarity 100.0%; Pred. No. 8.5e-122;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTQVSNVDVNLNDGASKYLNALAWTFNDSPNYYKTLGTQITPALFPKAGDILYSKDEL 61  
Db 45 QTQVSNVDVNLNDGASKYLNALAWTFNDSPNYYKTLGTQITPALFPKAGDILYSKDEL 104

Qy 62 GRTRTARGTLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSYVGDFWNR 121  
Db 105 GRTRTARGTLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSYVGDFWNR 164

Qy 122 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGILYYEVAP 181  
Db 165 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGILYYEVAP 224

Qy 182 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 229  
Db 225 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 272

RESULT 9  
US-10-453-032-15  
; Sequence 15, Application US/10453032  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; APPLICANT: Pang, Patty P.-Y.  
; APPLICANT: Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/453,032  
; FILING DATE: 03-JUNE-2003  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: 23-JUNE-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-10-453-032-15

Query Match 99.6%; Score 1215; DB 30; Length 272;  
Best Local Similarity 100.0%; Pred. No. 8.5e-122;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTQVSNVDVNLNDGASKYLNALAWTFNDSPNYYKTLGTQITPALFPKAGDILYSKDEL 61  
Db 45 QTQVSNVDVNLNDGASKYLNALAWTFNDSPNYYKTLGTQITPALFPKAGDILYSKDEL 104

Qy 62 GRTRTARGTLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSYVGDFWNR 121  
Db 105 GRTRTARGTLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVYKIEWLNGLSYVGDFWNR 164

Qy 122 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGILYYEVAP 181  
Db 165 SHLIADSLGGDALRVNAVGTGRTQNVGGRDQKGMRYTEQRAQEWLEARNRDGILYYEVAP 224

Qy 182 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 229  
Db 225 IYNADELIPRAVVVSMOSSDNTINEKVLVYNTANGYTIYHNGTPTQK 272

RESULT 10  
US-08-188-721A-2  
; Sequence 2, Application US/08188721A  
; GENERAL INFORMATION:  
; APPLICANT: Yutsudo, Takashi  
; APPLICANT: Okumura, Koichi  
; APPLICANT: Iwasaki, Makoto  
; APPLICANT: Hara, Ayako

```
; APPLICANT: Kishishita, Masamichi
; APPLICANT: Takeda, Yoshifumi
; APPLICANT: Igarashi, Hisanaga
; APPLICANT: Hinuma, Yorio
; TITLE OF INVENTION: Mitogenic Factor, Gene Thereof and
; TITLE OF INVENTION: Method of Microdetection Therefor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,721A
; FILING DATE: 31-JAN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1422-178P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 271 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-188-721A-2

Query Match 99.4%; Score 1213; DB 5; Length 271;
Best Local Similarity 99.1%; Pred. No. 1.4e-121;
Matches 227; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RQTQVSDVNLNDGASKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 60
Db 43 RQTQVSDVNLNDGATKYLNEALAWTFNDSPNYKTLGTSQITPALFPKAGDILYSLKDE 102
Qy 61 LGRTRTARGLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVKYKIEWNLGLSYVGDFTN 120
Db 103 LGRTRTARGLTYANVEGSYGVRSFGKNQNPAGWTGNPNHVKYKIEWNLGLSYVGDFTN 162
Qy 121 RSHLIADSLGDLARVNAVTRTONVGRDQKGMRYTEORAEWLEANDGYLYVEVA 180
Db 163 RSHLIADSLGDLARVNAVTRTONVGRDQKGMRYTEORAEWLEANDGYLYVEVA 222
Qy 181 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYINHGHTPTQK 229
Db 223 PIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYINHGHTPTQK 271

RESULT 11
US-10-360-101-263
; Sequence 263, Application US/10360101
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; TITLE OF INVENTION: Export and modification of (poly)peptide in the lantibiotic way
; FILE REFERENCE: 2183-5673
; CURRENT APPLICATION NUMBER: US/10/360,101
; CURRENT FILING DATE: 2003-02-07
; PRIOR APPLICATION NUMBER: EP 0207060.8
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 309
```

```
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 263
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: sequence of streptodornase
US-10-360-101-263

Query Match 29.9%; Score 365; DB 29; Length 303;
Best Local Similarity 34.6%; Pred. No. 5.9e-30;
Matches 84; Conservative 34; Mismatches 79; Indels 46; Gaps 4;

Qy 26 TFNDSPNYKTLGTSQITPALFPK-----AGDILYSKLDLGRTR 65
Db 19 TYGEKYDYTVIGESNIDQSAFKLYKTTTERRYVYKQGTSEKRVTVSDVNVNPLDGYKRST 78
Qy 66 TARTGLTYANVEGSYGVRSFGKNQNPAGW-----TGN-----PNH 101
Db 79 GAYGVTKDMIDMSKGYREKWTNPESGWFVFNADNEEISEKEYDSRRTKSKYKVTNN 138
Qy 102 VKYKIEWNLGLSYVGDFTNPSHLLIADSLGDLARVNAVTRTONVGRDQKGMRYTEQ 161
Db 139 VPVVLTTLTKGKYNHSLFVASHLFDLSLGSKSRKNAITGTQMNQVGR--KGGMQYIEK 196
Qy 162 RAQEWLEANDGYLYVEVAPIYNADELIPRAVVVSMOSSDNTINEKLVYNTANGTYIN 221
Db 197 KVLSHITKNDVYVFSAIPEYQGAELLARSLVLSALSSDGVINETVRVNTADGCFNIN 256
Qy 222 HNG 224
Db 257 EKG 259

RESULT 12
US-08-482-785-6
; Sequence 6, Application US/08482785
; GENERAL INFORMATION:
; APPLICANT: Adams, Craig W.
; APPLICANT: Pang, Patty P.-Y.
; APPLICANT: Belei, Marina
; TITLE OF INVENTION: Recombinant DNase B Derived from
; TITLE OF INVENTION: Streptococcus pyogenes
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheldon & Mak
; STREET: 225 South Lake Avenue, Ninth Floor
; CITY: Pasadena
; STATE: California
; COUNTRY: USA
; ZIP: 91001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,785
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/393,889
; FILING DATE: 24-FEB-1995
; APPLICATION NUMBER: US/08/082,845
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farber, Michael B.
; REGISTRATION NUMBER: 32,612
; REFERENCE/DOCKET NUMBER: 9521
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (818) 796-4000
; TELEFAX: (818) 795-6321
; INFORMATION FOR SEQ ID NO: 6:
```

SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-08-482-785-6

Query Match 16.6%; Score 202; DB 8; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38  
Db 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38

RESULT 13  
US-09-119-900-6  
Sequence 6, Application US/09119900  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/119,900  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-09-119-900-6

Query Match 16.6%; Score 202; DB 15; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38

Db 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38

RESULT 14  
US-10-453-032-6  
Sequence 6, Application US/10453032  
GENERAL INFORMATION:  
APPLICANT: Adams, Craig W.  
APPLICANT: Pang, Patty P.-Y.  
APPLICANT: Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
TITLE OF INVENTION: Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/453,032  
FILING DATE: 03-JUNE-2003  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: 23-JUNE-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 38 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
US-10-453-032-6

Query Match 16.6%; Score 202; DB 30; Length 38;  
Best Local Similarity 100.0%; Pred. No. 1.1e-13;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38  
Db 1 RQTQSVNDVVLNDGASKYLNEALAWTFNDSPNYYKTIG 38

RESULT 15  
US-09-689-278-2  
Sequence 2, Application US/09689278  
GENERAL INFORMATION:  
APPLICANT: Broudy, Thomas B.  
APPLICANT: Pancholi, Vijaykumar  
APPLICANT: Fischetti, Vincent A.  
TITLE OF INVENTION: Streptococcal Genes and Gene Products  
TITLE OF INVENTION: Induced by a Host Cell Contact  
FILE REFERENCE: 600-1-261N  
CURRENT APPLICATION NUMBER: US/09/689,278  
CURRENT FILING DATE: 2000-10-12



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: January 5, 2004, 18:39:19 ; Search time 36.2689 Seconds  
(without alignments)  
408.499 Million cell updates/sec

Title: US-08-482-785-9

Perfect score: 1220

Sequence: 1 RQGVNDVNDGASKYLN.....VYNTANGYTYNHNPTQK 229

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 320769 seqs, 64697744 residues

Total number of hits satisfying chosen parameters: 320769

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA New:\*

- 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*
- 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1220	100.0	229	4	US-08-472-630-9
2	1220	100.0	271	4	US-08-472-630-8
3	1215	99.6	272	4	US-08-472-630-15
4	202	16.6	38	4	US-08-472-630-6
5	170	13.9	32	4	US-08-472-630-16
6	122.5	10.0	274	1	PCT-US03-27401-436
7	122.5	10.0	274	6	US-10-472-928-4092
8	106	8.7	23	4	US-08-472-630-4
9	97	8.0	281	6	US-10-425-114A-51758
10	90	7.4	398	5	US-09-897-516A-6933
11	87.5	7.2	608	6	US-10-679-063-21735
12	87.5	7.2	1738	1	PCT-US03-28227-4134
13	87	7.1	428	1	PCT-US02-34769-8
14	87	7.1	429	1	PCT-US02-34769-16
15	86.5	7.1	331	6	US-10-425-114A-53765
16	86.5	7.1	446	6	US-10-425-114A-57900
17	86.5	7.1	450	6	US-10-425-114A-56379
18	86.5	7.1	1809	7	US-60-490-830-1987
19	86	7.0	270	6	US-10-425-114A-52835
20	86	7.0	329	6	US-10-425-114A-69718
21	86	7.0	396	6	US-10-679-063-26323
22	86	7.0	566	6	US-10-425-114A-47841
23	85.5	7.0	482	5	US-09-897-516A-6199
24	85.5	7.0	1475	1	PCT-US03-06362-34
25	84	6.9	362	6	US-10-250-682-2
26	83	6.8	1203	6	US-10-472-928-4086

27	83	6.8	1436	6	US-10-687-046-13	Sequence 13, Appl
28	82.5	6.8	818	5	US-09-897-516A-5123	Sequence 5123, Ap
29	82.5	6.8	1795	4	US-08-973-363A-17	Sequence 17, Appl
30	82	6.7	1639	6	US-10-679-063-17850	Sequence 17850, A
31	81.5	6.7	540	5	US-09-674-546A-2941	Sequence 2941, Ap
32	81	6.6	289	6	US-10-425-114A-44550	Sequence 44550, A
33	81	6.6	473	1	PCT-US03-12556-98	Sequence 98, Appl
34	81	6.6	811	6	US-10-272-998-7	Sequence 7, Appl
35	81	6.6	811	6	US-10-729-122-7	Sequence 7, Appl
36	81	6.6	945	5	US-09-614-150A-22485	Sequence 22485, A
37	81	6.6	2710	6	US-10-272-898-6	Sequence 6, Appl
38	81	6.6	2710	6	US-10-729-122-6	Sequence 6, Appl
39	80.5	6.6	285	6	US-10-009-384-30	Sequence 30, Appl
40	80	6.6	646	1	PCT-US03-27401-377	Sequence 377, App
41	80	6.6	646	6	US-10-472-928-2716	Sequence 2716, Ap
42	80	6.6	646	6	US-10-474-776-328	Sequence 328, App
43	79.5	6.5	750	5	US-09-581-286A-477	Sequence 477, App
44	79.5	6.5	819	5	US-09-581-286A-344	Sequence 344, App
45	79.5	6.5	843	1	PCT-US03-38193-1797	Sequence 1797, Ap

## ALIGNMENTS

### RESULT 1

US-08-472-630-9

; Sequence 9, Application US/08472630

; GENERAL INFORMATION:

; APPLICANT: Adams, Craig W.

; Pang, Patty P.-Y.

; Belci, Marina

; TITLE OF INVENTION: Recombinant DNase B Derived from

; Streptococcus pyogenes

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheldon & Mak

; STREET: 225 South Lake Avenue, Ninth Floor

; CITY: Pasadena

; STATE: California

; COUNTRY: USA

; ZIP: 91001

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/472,630

; FILING DATE: 07-Jun-1995

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/082,845

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Farber, Michael B.

; REGISTRATION NUMBER: 32,612

; REFERENCE/DOCKET NUMBER: 9521

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (818) 796-4000

; TELEFAX: (818) 795-6321

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 229 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Streptococcus pyogenes

; SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-08-472-630-9

Query Match

100.0%; Score 1220; DB 4; Length 229;

Best Local Similarity 100.0%; Pred. No. 2.4e-96; Indels 0; Gaps 0;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDE 60  
DB 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDE 60

QY 61 LGRTTRAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 120  
DB 61 LGRTTRAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 120

QY 121 RSHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 180  
DB 121 RSHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 180

QY 181 PIYNADLIPRAVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229  
DB 181 PIYNADLIPRAVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229

RESULT 2  
US-08-472-630-8  
; Sequence 8, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 271 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-08-472-630-8

Query Match 100.0%; Score 1220; DB 4; Length 271;  
Best Local Similarity 100.0%; Pred. No. 2.8e-96;  
Matches 229; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDE 60  
DB 43 RQTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDE 102

QY 61 LGRTTRAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 120  
DB 103 LGRTTRAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 162

QY 121 RSHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 180  
DB 163 RSHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 222

QY 181 PIYNADLIPRAVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 229  
DB 223 PIYNADLIPRAVVSMQSSDNTINEKVLVNTANGTYTINYHNGTPTQK 271

RESULT 3  
US-08-472-630-15  
; Sequence 15, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 272 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:

US-08-472-630-15

Query Match 99.6%; Score 1215; DB 4; Length 272;  
Best Local Similarity 100.0%; Pred. No. 7.6e-96;  
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDEL 61  
DB 45 QTVSNDVVLNDGASKYLNEALAWTFNDSPNYYKTLGTSQITPALFPKAGDILYSKLDEL 104

QY 62 GRTRTAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 121  
DB 105 GRTRTAGTTLTYANVEGSGYVRSFGKNQNPAGWGNPNHVKYKIEWLNGLSYVGDFWN 164

QY 122 SHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 181  
DB 165 SHLIADSLGGDALRVNAVGTGRTQNVGRDQKGGMRYTEQRAQEWLEARNRDLGYYEVA 224

Qy 182 IYNADELIPRAVVVSSOONTINEKVLVNTANGTYTINHGTPQK 229  
Db 225 IYNADELIPRAVVVSSOONTINEKVLVNTANGTYTINHGTPQK 272

## RESULT 4

US-08-472-630-6  
; Sequence 6, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belai, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 38 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-08-472-630-6

Query Match 16.6%; Score 202; DB 4; Length 38;  
Best Local Similarity 100.0%; Pred. No. 8.7e-11;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ROTQVSNVDVNLGASKYLNEALAWTFNDSPNYKTLG 38  
Db 1 ROTQVSNVDVNLGASKYLNEALAWTFNDSPNYKTLG 38

## RESULT 5

US-08-472-630-16  
; Sequence 16, Application US/08472630  
; GENERAL INFORMATION:  
; APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belai, Marina  
; TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes

; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheldon & Mak  
; STREET: 225 South Lake Avenue, Ninth Floor  
; CITY: Pasadena  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/472,630  
; FILING DATE: 07-Jun-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/082,845  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Farber, Michael B.  
; REGISTRATION NUMBER: 32,612  
; REFERENCE/DOCKET NUMBER: 9521  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (818) 796-4000  
; TELEFAX: (818) 795-6321  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptococcus pyogenes  
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-08-472-630-16

Query Match 13.9%; Score 170; DB 4; Length 32;  
Best Local Similarity 100.0%; Pred. No. 3.8e-08;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QTQVSNVDVNLGASKYLNEALAWTFNDSPNY 33  
Db 1 QTQVSNVDVNLGASKYLNEALAWTFNDSPNY 32

## RESULT 6

PCT-US03-27401-436  
; Sequence 436, Application PC/TUS0327401  
; GENERAL INFORMATION:  
; APPLICANT: TUFTS UNIVERSITY  
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSIS, TREATMENT AND  
; PREVENTION OF ACTIVE INFECTION  
; FILE REFERENCE: 700355-52941-PCT  
; CURRENT APPLICATION NUMBER: PCT/US03/27401  
; CURRENT FILING DATE: 2003-09-02  
; PRIOR APPLICATION NUMBER: US 60/407,082  
; PRIOR FILING DATE: 2002-08-30  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 436  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
PCT-US03-27401-436

Query Match 10.0%; Score 122.5; DB 1; Length 274;  
Best Local Similarity 22.3%; Pred. No. 0.0037;  
Matches 57; Conservative 34; Mismatches 100; Indels 65; Gaps 11;

Qy 3 QTQVSNVDVNLGASKYLNEALAWTFNDSPNYKTLGTSQI 42

Db 50 SQALESVLTDAVKSQIKGLEWNGSGAFIVNGKNTNLDAKSSKPYADNKTIVGKETV 109  
Qy 43 TPALFPKAGDILYSKLEDE--LGRTRTARTLTLYANVEGSYGVRSFGKQNPAGWTGNPN 100  
Db 110 -----PTVANALLSKATQYKRNKKTGNGSTSW-----TPPGW----- 142  
Qy 101 HVKYKIEWLNGLSYGVDFWNRSHLIADSL-----GGDALR-----VNAVGTGRTQNVGGRDQ 152  
Db 143 ---HQVKNLKG-SYTHAV-DRGHLGLYALIGLDGDFDASTSNPKNIIVQTAWANQAQAEY 197  
Qy 153 KGMRYTEORAEWLEARNRDGYLYEVAPIYNADE-LIPRAVVVSMQSSDNTINEKVLVY 211  
Db 198 STQNYYESKVRKALDQNK--RVRYRVTLTYASNEDLVPSAQIEAKSSDGELEFNVLVP 255  
Qy 212 NTANGYTYNHNGTPT 227  
Db 256 NVQKGLQDLYRTGEVT 271

RESULT 7  
US-10-472-928-4092  
; Sequence 4092, Application US/10472928  
; GENERAL INFORMATION:  
; APPLICANT: CHIRON SPA  
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH  
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS  
; FILE REFERENCE: P026926W0  
; CURRENT APPLICATION NUMBER: US/10/472,928  
; CURRENT FILING DATE: 2003-09-26  
; PRIOR APPLICATION NUMBER: GB-0107658.7  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 4979  
; SOFTWARE: Seqwin99, version 1.03  
; SEQ ID NO 4092  
; LENGTH: 274  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
; FEATURE:  
; OTHER INFORMATION: DNA-entry nuclease (enda)  
; OTHER INFORMATION: Cellular location: outside  
; OTHER INFORMATION: Similar to strain R6 sequence 15903821 (e-156)  
US-10-472-928-4092

Query Match 10.0%; Score 122.5; DB 6; Length 274;  
Best Local Similarity 22.3%; Pred. No. 0.0037;  
Matches 57; Conservative 34; Mismatches 100; Indels 65; Gaps 11;  
Qy 3 TQVSNVDVLDGASKYLNEALAWTFND-----SPNYKYTLGTSGI 42  
Db 50 SQALESVLTDAVKSQIKGLEWNGSGAFIVNGKNTNLDAKSSKPYADNKTIVGKETV 109  
Qy 43 TPALFPKAGDILYSKLEDE--LGRTRTARTLTLYANVEGSYGVRSFGKQNPAGWTGNPN 100  
Db 110 -----PTVANALLSKATQYKRNKKTGNGSTSW-----TPPGW----- 142  
Qy 101 HVKYKIEWLNGLSYGVDFWNRSHLIADSL-----GGDALR-----VNAVGTGRTQNVGGRDQ 152  
Db 143 ---HQVKNLKG-SYTHAV-DRGHLGLYALIGLDGDFDASTSNPKNIIVQTAWANQAQAEY 197  
Qy 153 KGMRYTEORAEWLEARNRDGYLYEVAPIYNADE-LIPRAVVVSMQSSDNTINEKVLVY 211  
Db 198 STQNYYESKVRKALDQNK--RVRYRVTLTYASNEDLVPSAQIEAKSSDGELEFNVLVP 255  
Qy 212 NTANGYTYNHNGTPT 227  
Db 256 NVQKGLQDLYRTGEVT 271

RESULT 8  
US-08-472-630-4  
; Sequence 4, Application US/08472630  
; GENERAL INFORMATION:

APPLICANT: Adams, Craig W.  
; Pang, Patty P.-Y.  
; Belei, Marina  
TITLE OF INVENTION: Recombinant DNase B Derived from  
; Streptococcus pyogenes  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheldon & Mak  
STREET: 225 South Lake Avenue, Ninth Floor  
CITY: Pasadena  
STATE: California  
COUNTRY: USA  
ZIP: 91001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,630  
FILING DATE: 07-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/082,845  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Farber, Michael B.  
REGISTRATION NUMBER: 32,612  
REFERENCE/DOCKET NUMBER: 9521  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (818) 796-4000  
TELEFAX: (818) 795-6321  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Streptococcus pyogenes  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-08-472-630-4

Query Match 8.7%; Score 106; DB 4; Length 23;  
Best Local Similarity 95.7%; Pred. No. 0.0072;  
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 QTVSNVDVLDGASKYLNEALA 24  
Db 1 QTVSNVDVLDGASKYLNEALA 23  
RESULT 9  
US-10-425-114A-51758  
; Sequence 51758, Application US/10425114A  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114A  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 51758  
; LENGTH: 281  
; TYPE: PRT







```
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73033D04_FLI.pep
US-10-425-114A-53765

Query Match      7.1%; Score 86.5; DB 6; Length 331;
Best Local Similarity 20.1%; Pred. No. 5.1;
Matches 58; Conservative 36; Mismatches 88; Indels 107; Gaps 13;

Qy 8 DVLNDCASK-----YLNEALAWTFNDSPNYYKTLGTQITPALFPKAGDILYSKLDELG- 62
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 2 DALVNTGLAKLVYEVVNIDDCWESDRDYQGSFVANRQTFFSGIKALADYVHAKGLKGI 61

Qy 63 ----RRTTAR-----GTLTYANVE-----GSYGV-----RQS 85
   ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 62 YSDAGTRTCSQKMPGSLDHEEQDVKTFFSSWGIDYLVKNDNCNDAGRSVMERVTKMSNAKMT 121

Qy 86 FGK-----NONPAGWTGNPNHVKYKIEWLNGLSYVQDFWNRSHLIADSLGGDALR 135
   : || : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 122 YKGIFPSLCEWGRQNPATWAG-----MGNSWRTTDDIADNWG----- 160

Qy 136 VNAVVTGTRTON-----VGGRDQ-----KGMRYTEORAEWLEANRDGYLYEVA 180
   : : | | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 161 --SMTSRADQNDRWASVAGPGGWNDDPDMLEVNGNGMSEAYRS-----HFSIW 206

Qy 181 PIYNADLIPRAV-VVSMQSSDNTINEKVLVYNT-----ANGYTINYHNG 224
   : | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 207 ALAKAPLLIGCDVRAMSQQTWGLSNSEVIAVNQDSQGAQGGKYQSSNG 255
```

Search completed: January 5, 2004, 18:56:47  
Job time : 37.2689 secs

**This Page Blank (uspto)**

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**This Page Blank (uspto)**